12

VSC-TR-81-30

EVALUATION OF THE MARS SEISMIC EVENT DETECTOR

W. E. Farrell
J. Wang
R. C. Goff
C. B. Archambeau

SYSTEMS, SCIENCE AND SOFTWARE
P.O. Box 1620
La Jolla, California 92038

TECHNICAL REPORT

August 1980

Monitored By:

VELA Seismological Center
312 Montgomery Street
Alexandria, VA 22314

AFTAC Project Authorization No. VT/0712/B/PMP
ARPA Order No. 2551, Program Code No. 6H189
Effective Date of Contract: November 17, 1978
Contract Expiration Date: November 15, 1981
Amount of Contract: $1,816,437
Contract No. F08606-79-C-0008
Principal Investigator and Phone No.
Dr. J. Theodore Cherry, (714) 453-0060
Project Scientist and Phone No.
Mr. Brian W. Barker, (202) 325-7581

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by AFTAC/VSC, Patrick Air Force Base, Florida 32925, under Contract No. F08606-79-C-0008.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER VSC-TR-81-30	2. GOVT ACCESSION NO. AD A114876	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) EVALUATION OF THE MARS SEISMIC EVENT DETECTOR		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER SSS-R-81-4656
7. AUTHOR(s) W. E. Farrell, J. Wang, C. B. Archambeau and R. C. Goff		8. CONTRACT OR GRANT NUMBER(s) F08606-79-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS Systems, Science and Software P. O. Box 1620 La Jolla, California 92038		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Program Code No. 6H189 ARPA Order No. 2551
11. CONTROLLING OFFICE NAME AND ADDRESS VELA Seismological Center 312 Montgomery Street Alexandria, Virginia 22314		12. REPORT DATE August 1980 13. NUMBER OF PAGES 55
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/ DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release, Unlimited Distribution.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Systems, Science and Software (S3) Multiple Arrival Recognition System (MARS) was run on two detection test tapes supplied by the VELA Seismological Center (VSC). The tapes contain nearly 45 hours of phase-randomized seismic noise in which a total of 268 events were added at varying signal levels. Compared against an algorithm using "optimum filters" which was run at the VELA Seismic Center, the MARS detector achieved a 13 percent greater →

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT (continued)

detection probability (with no greater false alarm rate) on both test tapes. A subsequent analysis at VSC using fixed bandpass filters, however, yielded results comparable to those reported here. Running on a Univac 1100/81 computer, the algorithm operates about 40 times faster than real time. The majority of the time is consumed in Fast Fourier Transforms.

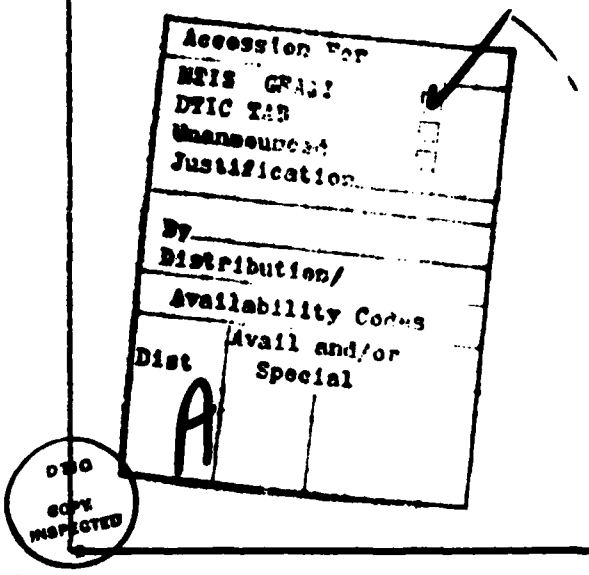

TABLE OF CONTENTS

Section			Page
	ABSTRACT		
I		SUMMARY AND RECOMMENDATIONS	1
II		MARS DETECTION ALGORITHM.	3
III		DETECTOR CALIBRATION	10
	3.1	CONTROL OF THE FALSE ALARM RATE.	11
	3.2	CONTROL OF THE PROBABILITY OF DETECTION.	16
	3.3	CHOOSING THE OPTIMUM DETECTOR.	17
IV		DETECTION TEST RESULTS	21
	4.1	NORSAR	22
	4.2	PINEDALE	31
	4.3	COMPARISONS WITH THE VSC BENCHMARK . . .	32
V		FUTURE MARS RESEARCH.	37
VI		CONCLUSIONS	39
		REFERENCES.	42
		APPENDIX 1 - EFFECT OF WINDOW LENGTH. . . . . ON FALSE ALARM RATE	43
		APPENDIX 2 - ESTIMATION OF SIGNAL-TO- . . . . NOISE RATIO FOR BROADBAND SEISMIC SIGNALS	44

LIST OF FIGURES

Figure		Page
1	Schematic representation of the three-dimensional parameter space spanned by the key MARS detection parameters.	4
2	MARS detector flow diagram.	5
3	Representation of the time-dependent amplitude. . . spectrum of a nonstationary seismic signal.	6
4	Variation in false alarm occurrence as the MARS . . event detection parameters are altered.	13
5	Variation in probability of detecting a signal as . the MARS event detection parameters are altered.	14
6	MARS detector receiver operating characteristic . .	18
7a	Detection results for the NORSAR test tape.	23
7b	. . . . .	24
7c	. . . . .	25
8a	Detection results for the Pinedale test tape. . . .	26
8b	. . . . .	27
8c	. . . . .	28
9	"Sonograms" for event 16 from the Pinedale test tape	38

LIST OF FIGURES

Figure		Page
A2.1	Probability of detecting a completely known signal, Q_d, plotted against Q_o the false alarm probability with parameter d the signal-to-noise ratio.	46
A2.2	Power spectrum calculated for a sample of Pinedale noise, with schematic representation of a typical signal spectrum.	47

I. SUMMARY AND RECOMMENDATIONS

A novel seismic event detector (MARS) based upon quasi-harmonic decomposition of broad-band time series has been tested on two data tapes constructed by the VELA Seismological Center. The principal results of the testing are:

- The MARS detector successfully finds significantly more weak events than does the STA/LTA algorithm with pre-filtering used at VSC;
- The enhanced probability of detection translates into a lower magnitude threshold at equal detection and false alarm probabilities. The MARS bodywave magnitude threshold is perhaps 0.1 magnitude unit lower than that of the current VSC detector;
- The Receiver Operating Characteristic (ROC) of the MARS detector has been defined for signal-to-noise ratios of about 2.0, and false alarm probability less than 0.02;
- Based upon the empirical ROC, the MARS detector is nearly as good as the optimum matched filter, but the inhomogeneous nature of the data set and the unresolved question of exactly how to relate false alarm rate to false alarm probability prevents exact quantification of this conclusion;
- The MARS detector seems to have a more regular false alarm behavior than the VSC detector, based on the two 24-hour data tapes;
- The MARS detector runs fast enough on a Univac 1100/81 that its implementation in a real time and on-line environment using a dedicated microcomputer is realistic;
- A procedure has been defined for calculating the signal-to-noise ratio for non-stationary seismic signals. This procedure, when applied to seismograms in the experimental data set, brings the empirical detection results in close accord with known results in statistical detection theory.

Based upon these results, the following recommendations are made:

- The experiment should be repeated using random number generators and spectral shaping programs to construct the "noise" data.

- The noise level was too stationary to adequately simulate typical seasonal variations in the noise spectrum. Greater noise variations should be used in the next set of experiments.

- The signals were too inhomogeneous, which makes signal-to-noise ratio calculations difficult. A fewer number of representative signal types should be used in the future.

- Radically new strategies should be developed to test regional phase (Lg) detectors.

- The MARS algorithm should be coded on a dedicated microcomputer and tested on real time data.

II. MARS DETECTION ALGORITHM

The Multiple Arrival Recognition System (MARS) seismic event detector searches for signals on the basis of three important characteristics: bandwidth, amplitude and dispersion (Figure 1). Each of these parameters may be adjusted to suit the observed noise and signal characteristics at a given station, thus allowing considerably more flexibility than a conventional STA/LTA power law detector. In addition, the inclusion of two detection criteria in addition to amplitude alone allows the amplitude threshold to be decreased so that weak signals have a higher detection probability.

The MARS detector uses multiple narrow band filters to decompose a seismic signal into its constituent frequency components. The instantaneous amplitude of each filter output yields the time-dependent power of the signal within the passband of the filter. The collection of amplitude functions from all the filters maps the signal onto the t_g-f plane, where t_g is group arrival time and f is frequency. The detector then uses an adaptive pattern recognition algorithm to search the t_g-f plane for undispersed peak alignments. Figure 2 displays the major elements of the algorithm in a flow diagram. Figure 3 displays a map of the t_g-f plane, and shows how information in this plane is related to both the original seismogram and its spectrum.

The filters used in MARS have a Gaussian frequency response and are generally of high-Q. They are equally spaced in frequency so as to "comb" the frequency band containing the signal. The spacing and Q are chosen so that adjacent filters yield signals that are relatively independent — i.e., their outputs for a random white noise input are less than 50 percent correlated. The filter Q's allow time and frequency resolution to be traded against one another within the fundamental governing relationship $\Delta\omega\ \Delta t \sim 1/2$.

The output of each filter is used to compute narrow band envelope functions proportional to the instantaneous signal amplitudes near the filter center frequencies. Each envelope

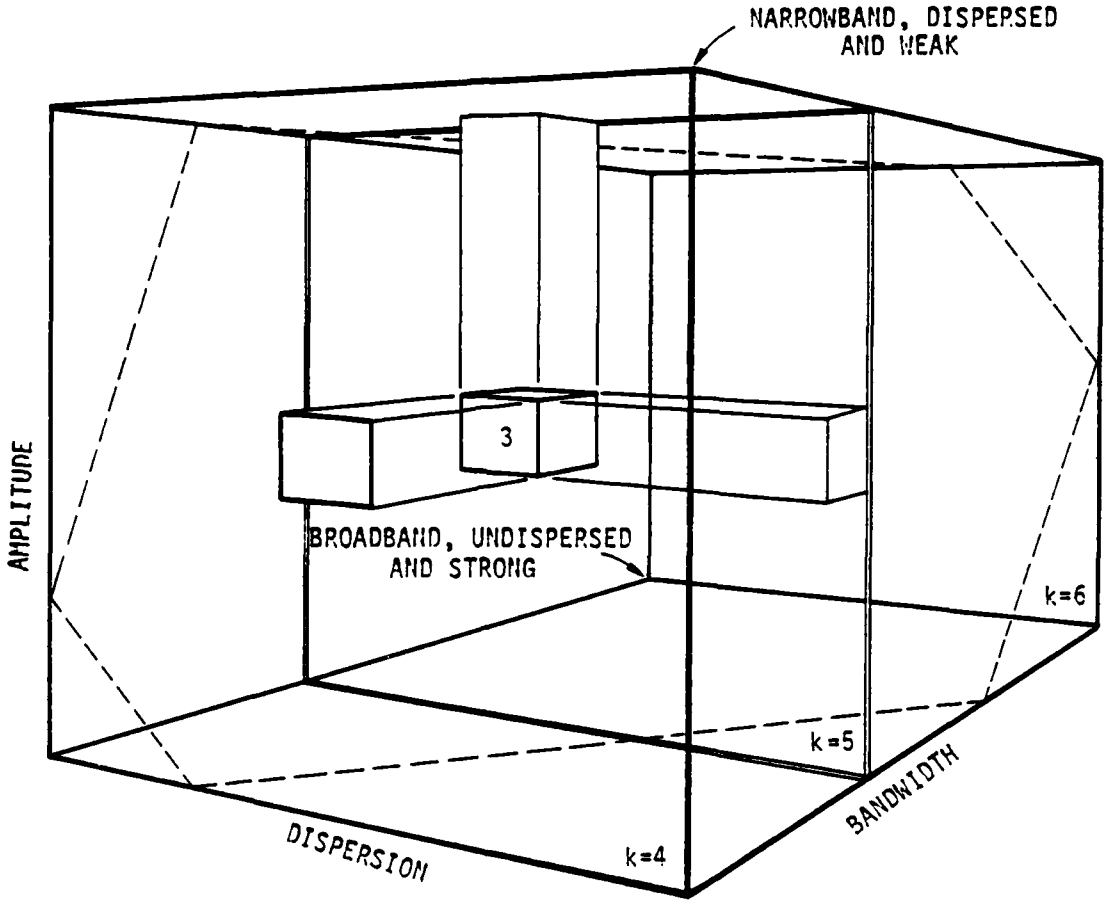

Figure 1. Schematic representation of the three-dimensional parameter space spanned by the key MARS (Multiple Arrival Recognition System) detection parameters. Each cell within this cube contains a number giving the false alarm rate when the algorithm operates on pure noise. The dashed lines across each face show the intersection of the cube with a nearby plane surface upon which the false alarm rate is constant.

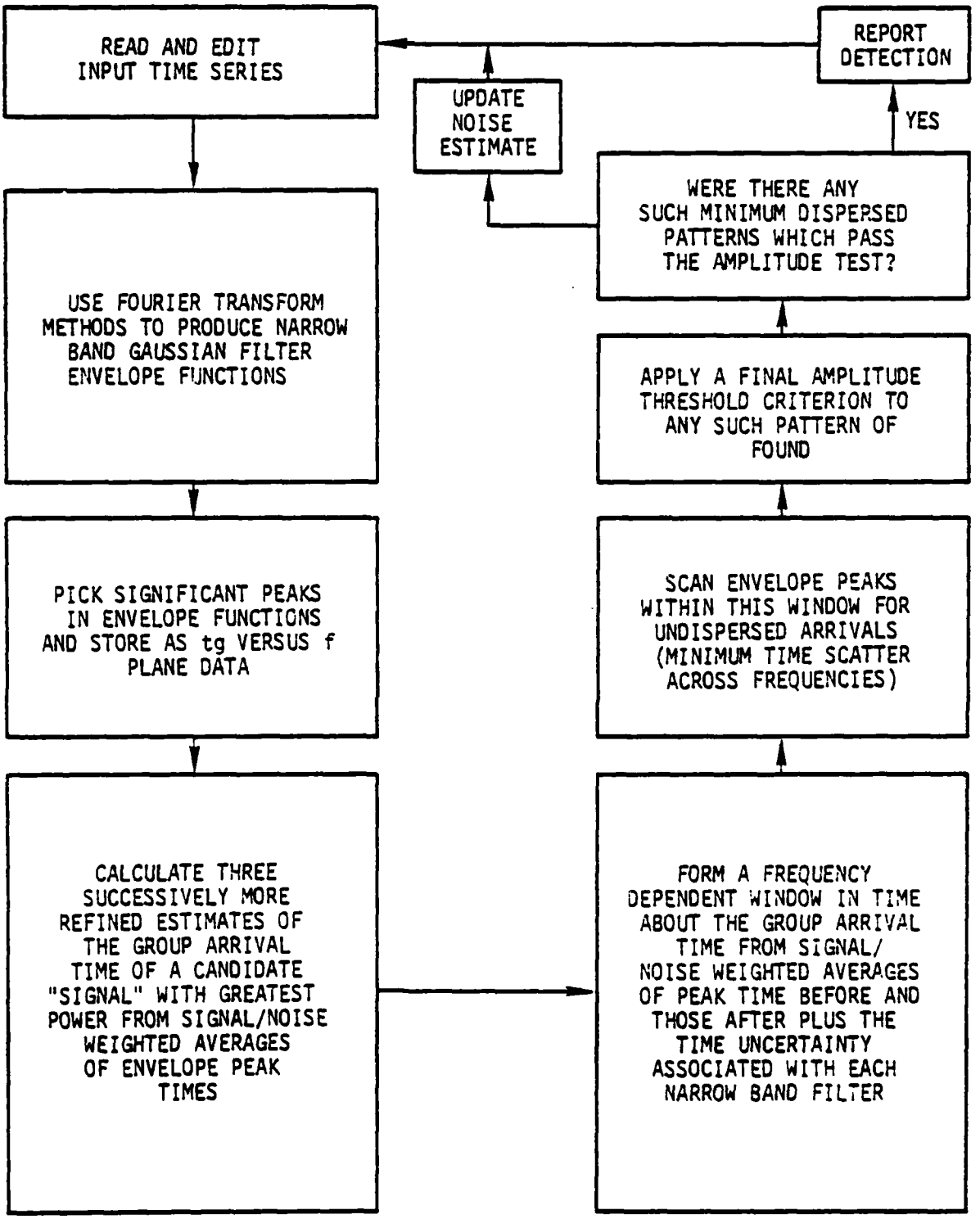

Figure 2. MARS detector flow diagram.

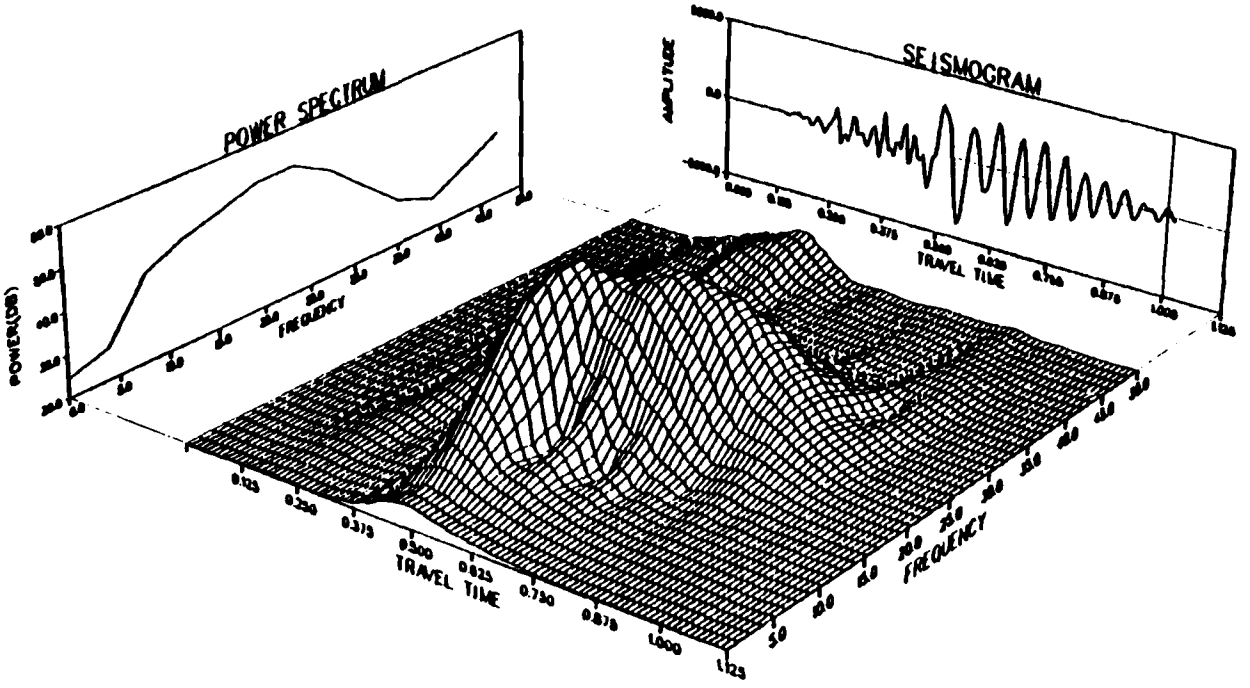

Figure 3. Representation of the time-dependent amplitude spectrum of a nonstationary seismic signal. The surface display shows the instantaneous signal amplitude as a function of time and frequency. When the (complex) amplitude is totalled over all frequencies at each time instant, the seismogram itself is recovered. When the (complex) amplitude at each frequency is totalled over all times, the power spectrum is obtained. These two one-dimensional data representations are shown projected back in the proper direction. The time series has the best possible time resolution, which is determined by the sample rate. The power (or Fourier) spectrum has the best possible frequency resolution, which is governed by the total duration of the data window. The power spectrum shown here, however, has been smoothed in frequency to show the overall trend more clearly. The two-dimensional surface plot of the time-dependent power (or amplitude) compromises both the time resolution of the seismogram itself and the frequency resolution inherent in the signal duration in order to map more clearly the time-dependent changes in the signal. There is an infinite variety of possible trade-offs between frequency resolution and time resolution, but they are connected by the uncertainty principal $\Delta f \Delta t = 1/4\pi$.

function is then scanned for local maxima, or peaks. The peak amplitudes and peak arrival times are written into an intermediate file before the pattern recognition algorithm is run.

Typically, envelope functions are generated from segments of data 1024 to 2048 points long. The segments themselves overlap in time by about ten percent, because the ends of each segment are tapered to reduce the effect of wrap-around in the Fourier transform filtering calculation. With sample rates of 10 to 20 Hz, the segments contain between 50 and 200 seconds of data. MARS then targets the time of the most probable "event" in each section by taking a weighted average of the times of the largest peak in each frequency band. The weighting gives more influence to peaks relatively higher above the mean noise amplitude in each band.

Undispersed alignments of peaks are sought within a wide search window about the target time. Events are then found by scanning the wide search window with a narrow window. Whenever a prescribed number of peaks fall into the narrow window, a "collection" is formed. An event or phase is declared if the collection contains a specified number of peaks above a preselected amplitude threshold level.

If no signal is detected, it is assumed that the time window contains nothing but noise. In this case the mean and standard deviation of the amplitudes of the peaks in each frequency band are computed and exponentially averaged with previous estimates. This running noise estimate is essentially a LTA-type average, but differs from conventional methods in that it adaptively conforms to changes in the shape and amplitude of the noise power spectrum. The high-level amplitude threshold is defined in terms of the number of standard deviations above the mean peak amplitude.

The significant parameters in MARS can be summarized as follows:

1. The dispersion parameter δ, is the width (in time) of the narrow window used in collection scanning.

2. The amplitude parameter, γ, defines the threshold amplitude which separates noise peaks from signal peaks.

3. The bandwidth parameter, k, gives the number of high-level peaks necessary to declare an event.

An event, whether true signal or false alarm, often produces more than one phase that passes all the currently implemented pattern recognition tests. This multiple phase behavior, however, is much more pronounced when the event actually is a signal, because most signals have a coda significantly longer in time than the narrow band filter impulse response. When multiple phases occur, they are ranked according to their mean relative amplitude (MRA). The MRA is simply the average amplitude of all peaks in a collection, with each peak weighted inversely by the mean noise amplitude in that particular frequency band.

The most important phases (sometimes as many as twenty are flagged) to consider are the earliest phase and the strongest (highest MRA) phase. With a short, undispersed P-wave arrival, the arrival time of the highest MRA phase is generally more closely aligned with the visible first motion than any other phase. The earliest phase picked for these types of signals is commonly 6 seconds earlier than the apparent first motion. On the other hand, regional seismic signals such as Lg can last for many tens of seconds and the highest MRA phase picked by the algorithm is often towards the middle of the record. For this type signal, the earliest detected phase is usually much closer in time to the true first motion.

This report discusses results obtained from analysis of two test tapes provided by the VELA Seismological Center. The first test tape contained about 23 hours of scrambled NORSAR noise to which were added numerous signals recorded at other times on the same, or similar, NORSAR instruments. The second tape was fabricated in an analogous fashion from records off the Pinedale SRO seismometer. In both tapes approximately thirty signals were added at specified times. Each signal was included at four different

amplitude levels, referred to as raw signal-to-noise ratio 1/2, 1/4, 1/8, 1/16. We call them the Class A through Class D signals.

Numerous enhancements to the MARS processor were implemented for this study. This has resulted in a code optimized for signal detection applications which is radically different from the earlier version used for seismic discrimination. In addition, the program was divided into two separate pieces to permit economical experimentation with a variety of detection parameter choices. In actual operation, of course, the pieces would be rejoined. This would significantly reduce the processing time, since a good fraction of the time is spent storing and retrieving data from the disc file.

The first independent program took the data, split it into segments, performed the narrow band filtering and created what is referred to as a t_g-f file, a catalogue for each segment of the times and amplitudes of the envelope maxima occurring in each filter output. Fourier transformations are the principal operation within this program, which processes 24 hours of data in about 30 minutes. The second program contains the detection algorithm. Several different parameter sets were used in the course of the study. Segmenting the code permits this iterative analysis with different parameter sets without requiring the recreation of the t_g-f data each time. The detection algorithm alone processed the filtered data in less than ten minutes.

III. DETECTOR CALIBRATION

The statistical theory of signal detection teaches tnat there is always a tradeoff between the probability of recognizing a real signal and the probability of mistakenly supposing one to occur, even if there is only random noise present. Tne crux of the problem is then simultaneously to maximize tne probability of detection and to minimize the false alarm rate (FAR).

Numerical experiments were performed to study the trade-off between detection probability and false alarm rate as a function of MARS DETECTION PARAMETERS. The principal result of these studies was the selection early on of a set of narrow band filter center frequencies and bandwidths. For Norsar, 20 filters equally spaced between .25 and 5.0 Hz were used: The Pinedale data was processed through 21 filters equally spaced between .5 and 3.0 Hz. These choices reflect the differences between the narrow band, low frequency Pinedale signals, and the higher frequency NORSAR signals. In all instances, filters of constant bandwidth were used. In terms of the filter Q, since $Q = f/\Delta f$, a constant Δf filter has a Q linearly related to the center frequency. Best results were usually obtained for $Q = 12\ f$.

The filter bandwidth choice is quite distinct from the signal bandwidth parameter appearing in Figure 1 and the related discussion of this section. In fact, the parameter space sketch shown in Figure 1 really applies for each distinct frequency comb through which the data is passed. Although we have not obtained definitive results on how the choice of comb affects detector performance, it is a much more tolerant parameter than the three principal parameters more completely analyzed.

The general philosophy in establishing the filter comb is that the time resolution (proportional to $1/\Delta f$) ought to be about the length of the signal time history, the frequency spacing ought to be about equal to the filter width, and the total frequency band

spanned should cover the frequency range of the expected signals. The choice of frequency spacing was based on analytical and numerical calculations of the coherency between adjacent envelope functions.

3.1 CONTROL OF THE FALSE ALARM RATE

The first step in the present detection experiment was to calibrate the algorithm's false alarm behavior. This was accomplished empirically by processing noise samples from each of the two seismic stations NORSAR and Pinedale several times. This exploratory study was guided by comparisons between event spectra and signal spectra, which targeted the frequency band with the highest signal-to-noise ratio (SNR). Typically the band was 0.8 Hz to 3.5 Hz. Furthermore, virtually all events are undispersed. Thus, the bandwidth parameter range and the dispersion parameter range are closely bracketed. These two attributes by which signals can be differentiated from random noise are most clearly elucidated by time-dependent power spectra such as those shown in Figure 3 and Figure 9.

For selecting a suitable amplitude range, statistical analyses were made of narrow band envelope functions and narrow band envelope peaks. These calculations yielded histogram approximations to the probability density and probability distribution function. To a good approximation, the probability density of continuous narrow band envelope functions and isolated envelope peaks follow the Rayleigh law, as is predicted for narrow band signals extracted from broad band Gaussian noise. The key result of the calculations was the identification of a rather narrow amplitude range over which the probability distribution jumped from 0.1 to 0.9. Put more simply, small amplitude envelope peaks are very common, large amplitude peaks quite rare and it is possible to quantify the transition region between these two parts of the probability distribution function. Within this transition region there are few enough peaks expected on chance alone that a useful false alarm rate might result (depending on how the dispersion and bandwidth parameters are chosen).

Having bracketed rather closely the interesting volume of the parameter space (see Figure 1), more detailed numerical experiments were made to explore the false alarm performance of the algorithm within this range. Figure 4 presents results obtained from multiple analyses of a 3.5 hour section of noise data extracted from the Pinedale test tape. Displayed in the figure are three matrices, each element of which represents a cube within the parameter space sketched in Figure 1. Thus, each cell represents a unique amplitude-dispersion-bandwidth triplet and hence a unique signal detection algorithm.

The most noticeable feature of the false alarm matrices is the sensitivity of the false alarm rate to the amplitude threshold. If M(f) denotes the mean noise amplitude in a frequency band and S(f) the corresponding standard deviation, then the threshold, T, is given by

$$T = M + \gamma S \qquad (2.1)$$

The parameter γ labels the left hand amplitude coordinate in both Figures 4 and 5. Only envelope peaks larger than T are allowed as candidate signal peaks.

Based on the observation that envelope peaks are approximately Rayleigh distributed, it follows from the Rayleigh density function (Papoulis, 1965, p 195) that S = 0.52 M. The threshold amplitude scale can then be converted to decibels relative to mean noise amplitude according to

$$dB = 20 \log (1 + 0.52\gamma) \qquad (2.2)$$

This coordinate labeling is given on the right hand axis of both Figures 4 and 5.

Reading up the columns, it is remarkable now much the false alarm rate varies for a 0.93 dB (10 percent) variation in the amplitude threshold.

A reasonable FAR target is perhaps five/hour, or 17 in this 3.5 hour noise record. It can be seen in Figure 4 that there are

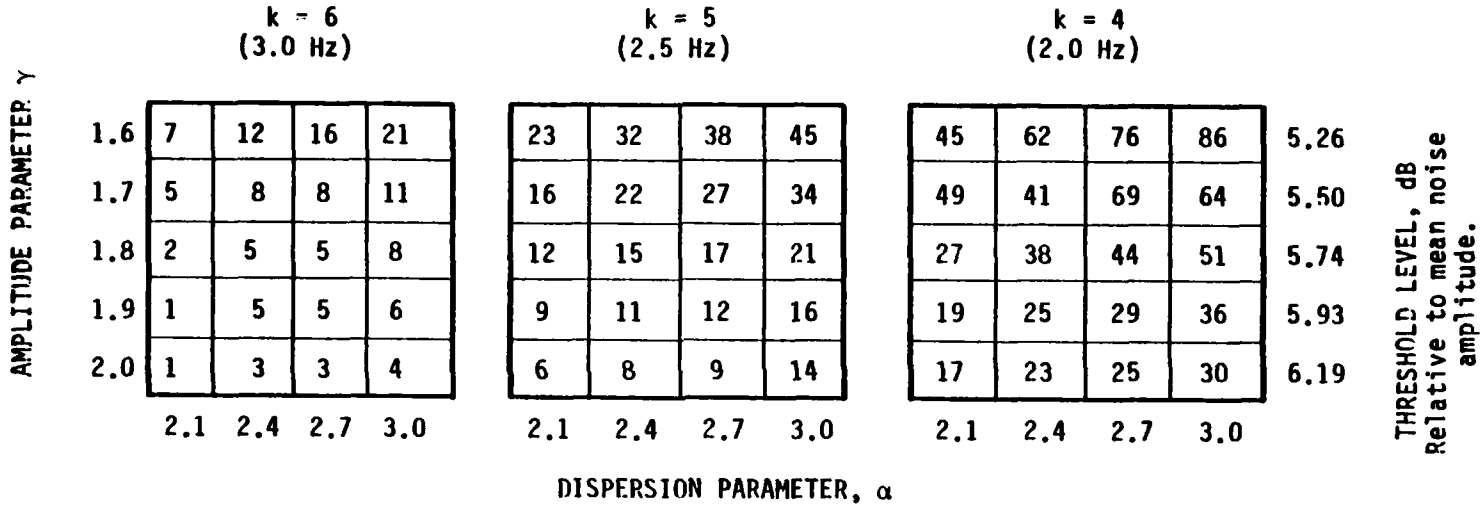

Figure 4. Variation in false alarm occurrence as the MARS event detection parameters are altered. Moving from left to right, there are three constant bandwidth planes extracted from the parameter space displayed in Figure 1. Within each plane, the vertical axis is proportional to the amplitude parameter and the horizontal axis is proportional to the dispersion parameter. The number contained in each cell gives the total number of false alarms which resulted from that particular parameter triplet when a 3.5 hour section of Pinedale noise was processed.

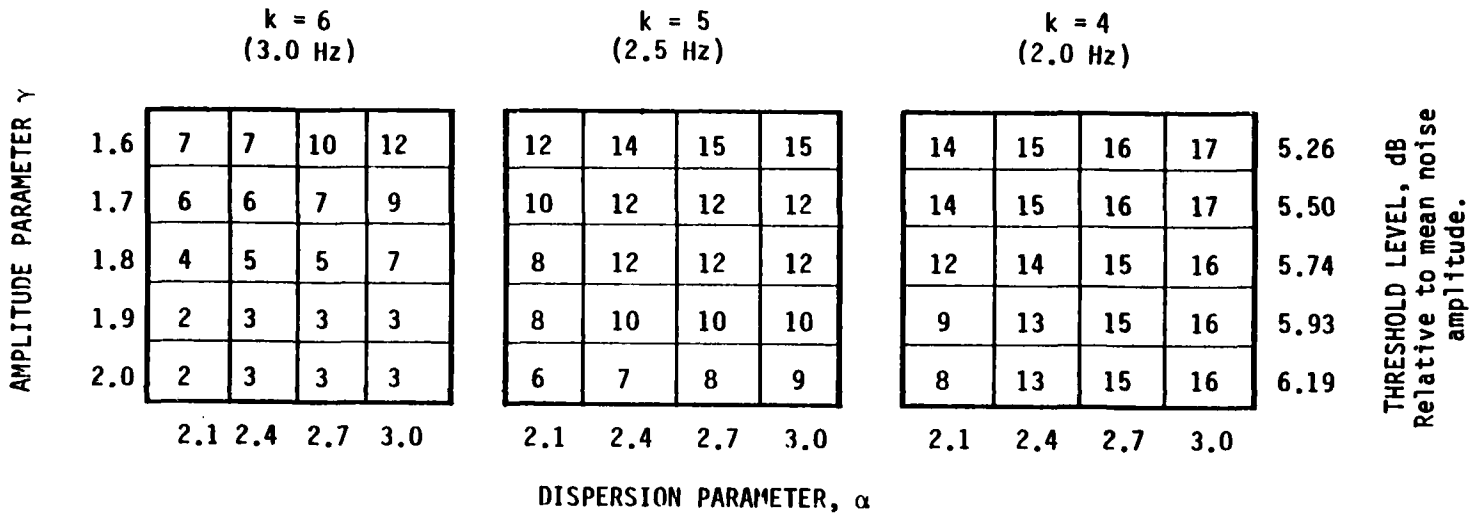

Figure 5. Variation in probability of detecting a signal as the MARS event detection parameters are altered. This figure is like Figure 4, except that the numbers in each matrix give the total number of Class B signals (out of 31 possible) successfully identified on the Pinedale test tape. Each numbered point on this figure and on Figure 4 represents a unique signal detection algorithm.

many detection parameter combinations which result in 17 false alarms. The dependence of FAR on detection parameters is surprisingly regular, and accords with intuition. Large, broad band, and undispersed signals are rare, while small, narrow band dispersed occurrences are common.

If one visualizes the three planes of Figure 4 slid into the parameter space cube of Figure 1 in the indicated order, the false alarm numbers can be thought of as defining a scalar field. The gradient of the field, the directions in which the false alarm rate changes most rapidly, tends to point from the lower lefthand corner of the back plane towards the upper righthand corner of the front plane. The level surfaces, which are perpendicular to the field gradient, define sheets on which the false alarm rate is constant. They tend to form planes cutting the front plane of the parameter space cube at an oblique angle near the bottom left corner, and intersecting the back plane towards the top right corner. The dashed lines on each cube face indicate how such a plane might be oriented. These exploratory calculations showed how to "tune" the parameters to control the false alarm rate.

A minor point to be mentioned is the relationship between the false alarm rate achieved by the VSC benchmark calculation and the false alarm rate achieved by the MARS calculation. They are not exactly equivalent, for the STA/LTA algorithm intrinsically operates on each individual data sample, whereas the MARS process takes a large block of data (75 seconds) and analyzes it as a unit. It thus can happen that in a given block, the STA/LTA procedure might pick two or more "events," but the MARS detector will only catch the largest of these. It is shown in Appendix 1 that for the low false alarm rates of interest, this interference effect is negligible and if the window length of the MARS detector were reduced by a factor of fifty, the false alarm rate would only rise by twelve percent.

3.2 CONTROL OF THE PROBABILITY OF DETECTION

Complementary to the false alarm behavior discussed previously is the analogous problem of quantifying how the detection parameters affect the probability of correctly identifying a signal known to be present. To study the probability of detection, the Pinedale Class B (raw SRN = 1/4) events were sought with the same variety of parameters used in the false alarm study. This numerical study was inevitably less precise than the FAR analysis because of the inhomogeneity and paucity of the signals. The Pinedale data was used because it seems more homogeneous than the NORSAR data. The Class B events were most suitable because both the optimum MARS results and the VSC benchmark results indicated that about half could be reliably picked out. Using twenty filters to span the frequency band from 0.5 Hz to 3.0 Hz, and choosing Q = 12 f (Δf = 0.08 Hz) multiple passes of the detection algorithm produced the results presented in Figure 5.

The tenor of this figure follows closely the false alarm data depicted in Figure 4. In both cases, the lower lefthand corner of the left panel has a *small number* (*low FAR implies small probability* of detection), while the top right corner of the right panel has a big number (nigh FAR implies large probability of detection). Thus, if these three planes are imagined slipped back into the parameter space cube, with the number of detections in each cell constituting a scalar field, then the gradients and the level surfaces are oriented much as before.

There is a notable distinction between the detection probability and the false alarm probability results. Comparing Figures 4 and 5 it is seen that the ordering of the numbers is similar, but the range quite different. In Figure 4, the greatest number of false alarms is nearly 100 times the least number. In Figure 5 the ratio of the greatest number of detections to least number of detections is less than ten even though the parameter sets are the same. This phenomenon is well known in the the theory of statistical signal detection and the result obtained here is not too different from the behavior of the matched filter.

Appendix 2 contains a discussion of the relation between detection probability, false alarm probability and signal-to-noise ratio. There is also an approximate calculation of the signal-to-noise ratio for the seismograms in the Pinedale tape. For the Class A events it is about 4.0, so for the Class B events the signal-to-noise ratio is about 2.0. Taking this value for the signal-to-noise ratio, we examine the trade off between detection probability and false alarm probability as depicted in Figure A2. The point labeled D shows 50 percent detection probability (15 out of 30). Dropping down to 5 percent probability but holding d constant, it can be estimated that the false alarm probability drops by about the factor of 100 seen in this numerical experiment.

3.3 CHOOSING THE OPTIMUM DETECTOR

It is clear that there is no "best" detector, that is, there exists no perfect combination of the amplitude-bandwidth-dispersion parameters. It is, however, not too difficult to select an optimum detector, for the phenomenon of diminishing returns, holds here with a vengeance. One must accept the fact that beyond some point, only modest increases in signal detectibility will be gained at the expense of order-of-magnitude increases in false alarm rate.

For any signal detector, the relationship between detection probability, Q_d, and false alarm probability, Q_0, at fixed signal-to-noise ratio is known as the Receiver Operating Characteristic. Using the false alarm counts (Figure 4) and the Class B detection counts (Figure 5) for the Pinedale data, we have constructed a receiver operating characteristic (Figure 6). To do this, the successful detection counts were classified by increasing probability. For each Q_d probability level the maximum and minimum false alarm counts were located for the matrices of Figure 4. These were turned into false alarm probability limits by the rule (Berger, 1980)

$$Q_0 = FAR \cdot T_d \qquad 3.1$$

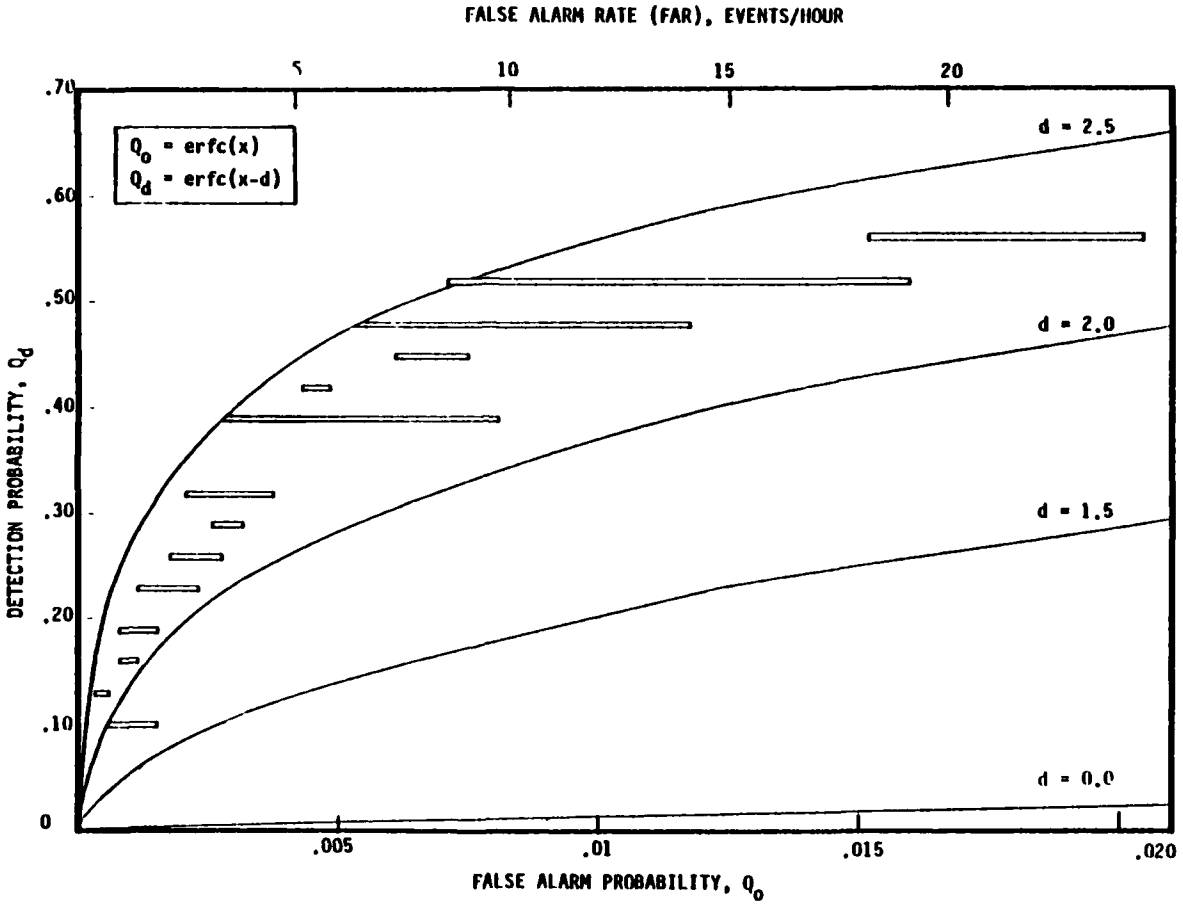

Figure 6. MARS detector receiver operating characteristic (ROC). The four curved lines show the ROC of the matched filter detector for signal-to-noise ratios, d, ranging from 0.0 to 2.5. These curves are defined by the parametric equations in the upper left box. The horizontal bars show the range of the ROC for the MARS detector derived from Figures 4 and 5. See the text for the method of construction.

The False Alarm Rate, FAR, is the number of false alarms divided by 3.5 hours, the length of the time series analyzes. The detection time, T_d, is the width (in hours) of the dispersion parameter. Thus

$$T_d = \alpha\ T_u/3600 \qquad 3.2$$

where α is the dispersion parameter and T_u the time uncertainty of the filter comb.

Using the approximate result that the Pinedale Class B signals have a signal-to-noise ratio of about 2, we also show in Figure 6 the ROC for the matched filter detector, and white noise spectrum, for signal-to-noise ratios of 0.0, 1.5, 2.0, 2.5.

Although the false alarm rate is a clearly understood notion, it is not so clear to us how false alarm rates are properly converted to false alarm probabilities. Clearly a time-like quantity enters as is expressed by equation 3.1 (Berger, 1980). We choose to make T_d depend on the dispersion parameter, so that, looking back at Figure 4, equal numbers in different columns translate into different Q_d values.

It is thus natural to follow the customary approach of fixing a priori a manageable false alarm rate, and then accepting whatever detection probability and signal-to-noise ratio combination follows from that. Even then the solution for the optimum detector is ambiguous. Looking back again at Figures 4 and 5, if one visualizes the surface on which the number of false alarms is 20 and number of detections is 12, they are not very far apart throughout an appreciable region of the parameter space.

The near tangency of the probability of detection surface and the probability of false alarm surface is a valuable attribute. It means that once a point (i.e., a set of detection parameters) has been chosen near the center of the region of tangency, then first order variations in parameter values lead only to second order variations in detector performance. If one invents a measure of

detector performance (one might be the ratio of detection probability to false alarm probability) defined on the vector field of detection parameters, then the measure has a broad flat maximum about the "best" parameter set. A statistical quantity with this property is said to be robust.

The parameter sets finally adopted for processing the two test tapes were slightly different, but roughly corresponded to the cell in the middle panel of Figures 4 and 5. For the NORSAR tape, for example, the following set was used.

T	Segment length	75 seconds
n	number of filters	20
f_c's	filter center frequencies	equally spaced every 0.25 Hz from 0.25 Hz to 5.0 Hz
Q's	filter Q's	Q = 12f for all filters.
δ	dispersion parameter	2.8 seconds
k	bandwidth parameter	4 peaks above threshold (total signal bandwidth at least 1 Hz)
γ	amplitude parameter	1.7 standard deviations above mean

The following different parameters were used in the Pinedale analyses.

n	number of filters	21
f_c's	filter center frequencies	equally spaced every .125 Hz from 0.5 Hz to 3.0 Hz
k	bandwidth parameter	5 peaks above threshold (total signal bandwidth at least .625 Hz)
γ	amplitude parameter	1.8 standard deviation above mean

IV. DETECTION TEST RESULTS

Each test tape contained approximately 24 hours of phase-scrambled noise in which more than 30 distinct events were included, each at four different gain levels. These gain levels are referred to in the VSC correspondence as raw signal-to-noise ratio = 1/2, 1/4, 1/8, 1/16. It is shown in Appendix 2 that by a more conventional definition of signal-to-noise ratio, the progression is roughly 4, 2, 1, 1/2 for the short duration teleseisms. We have found it more convenient simply to refer to the events as Class A (largest) through Class D, using the notation 2C, for instance, to connotate the next weakest copy of event 2.

The events were equally spaced in time, with the "arrival" coming at the start of the ninth minute of each ten-minute section of data. Thus, with segments 75 seconds long, events occur every eight segments. A MARS time pick was considered valid if it could be visually identified with a phase on the "A" level signals. (This criterion implies that visual identification is still the best event detector at large signal-to-noise ratios).

The tapes were processed in three steps. As described previously, short runs were targeted at selected events to explore the effect of detection parameters on false alarm rate and probability of detection. On the basis of these studies, the filter comb was established and the t_g-f file created. Using the t_g-f file as data, the detector algorithm was executed using the detector parameter combination that yielded the best detection probability at fixed FAR.

Results of the detection experiments are presented in two different formats. For the first format, analog plots were made of the Class A events. These were extracted from the continuous data tape, and filtered with a high pass filter to suppress the noisy band between zero frequency and 0.8 Hz. This makes the signals much more visible. On each signal, then, the times at which phases were detected were plotted, not only for the raw SNR = 1/2 segment (Class

A signals), but for the three weaker levels as well. These summary plots are presented as Figures 7 and 8. In addition, the results are presented in tabular form (Tables 1 and 3) with a layout similar to that in which the VSC benchmark results were presented. The VSC benchmark results are reproduced as Tables 2 and 4. Since a variety of signal types were included in the test tape, the former analog presentation makes clear the relationship between time domain appearance and detector performance.

4.1 NORSAR

Most of the events on the NORSAR tape are teleseismic with $4.0 < m_b < 6.0$ and with significant energy confined to the 1 to 3 Hz band. In most cases the Class A events were easily seen by eye, particularly when the data was high-passed through a filter with a 1 Hz corner frequency.

MARS declared a total of 158 "events" for the 23 hours of data. Of these, 60 occurred in the signal segments and 52 were accepted as valid, based upon visual comparison with phases in the Class A signals (Figure 7). Thus, 106 events were actually false alarms. Using the looser criterion of accepting any picks within 30 seconds of the nine-minute arrival time, there are 58 valid picks and only 100 false alarms. However, we feel that visually correlating time picks with the seismograms is a more definitive measure of the detector performance.

The MARS detector (Table 1) compares favorably with the optimally-filtered STA/LTA detector based upon the benchmark provided by VSC (Table 2). Out of the 33 Class A events, MARS missed seven. Of these, four (6, 19, 22 and 32) contain no visible signal. Events 4 and 15 are narrow band signals, and do not satisfy the MARS bandwidth criterion. Event 28 is a six-second dispersed arrival that exceeds the MARS dispersion parameter of 2.8 seconds.

Also detected were 16 Class B signals, 7 Class C signals, and 3 Class D signals. Included in the 52 total detections are 10

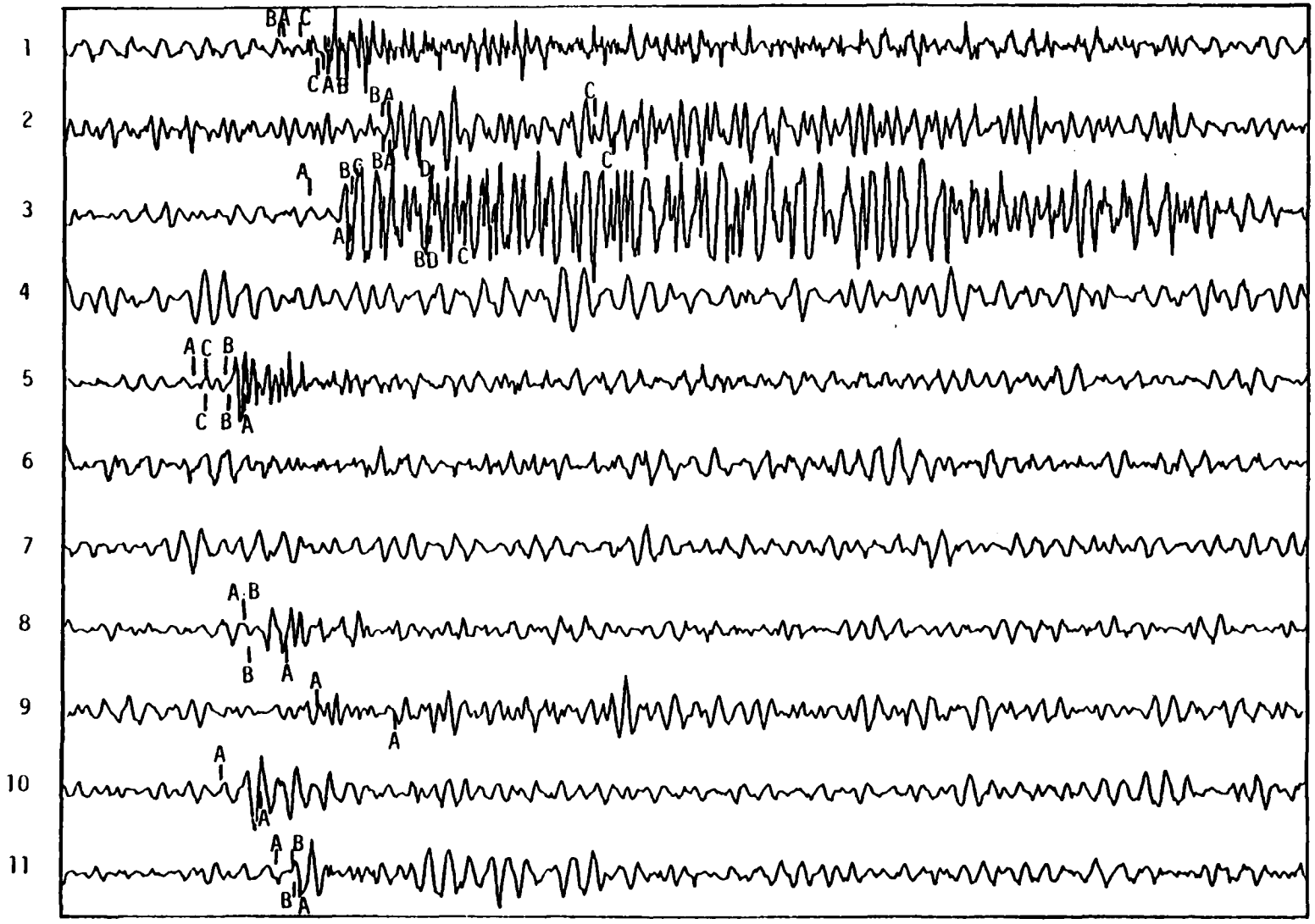

Figure 7a. Detection results for the NORSAR test tape. Letters A through D show the amplitude level of the signal for which the labeled phase was detected. Symbols above each trace apply to earliest picks; symbols below each trace apply to picks with the largest mean relative amplitude (MRA). Overall false alarm data, as well as a summary of detections, are presented in Table 1.

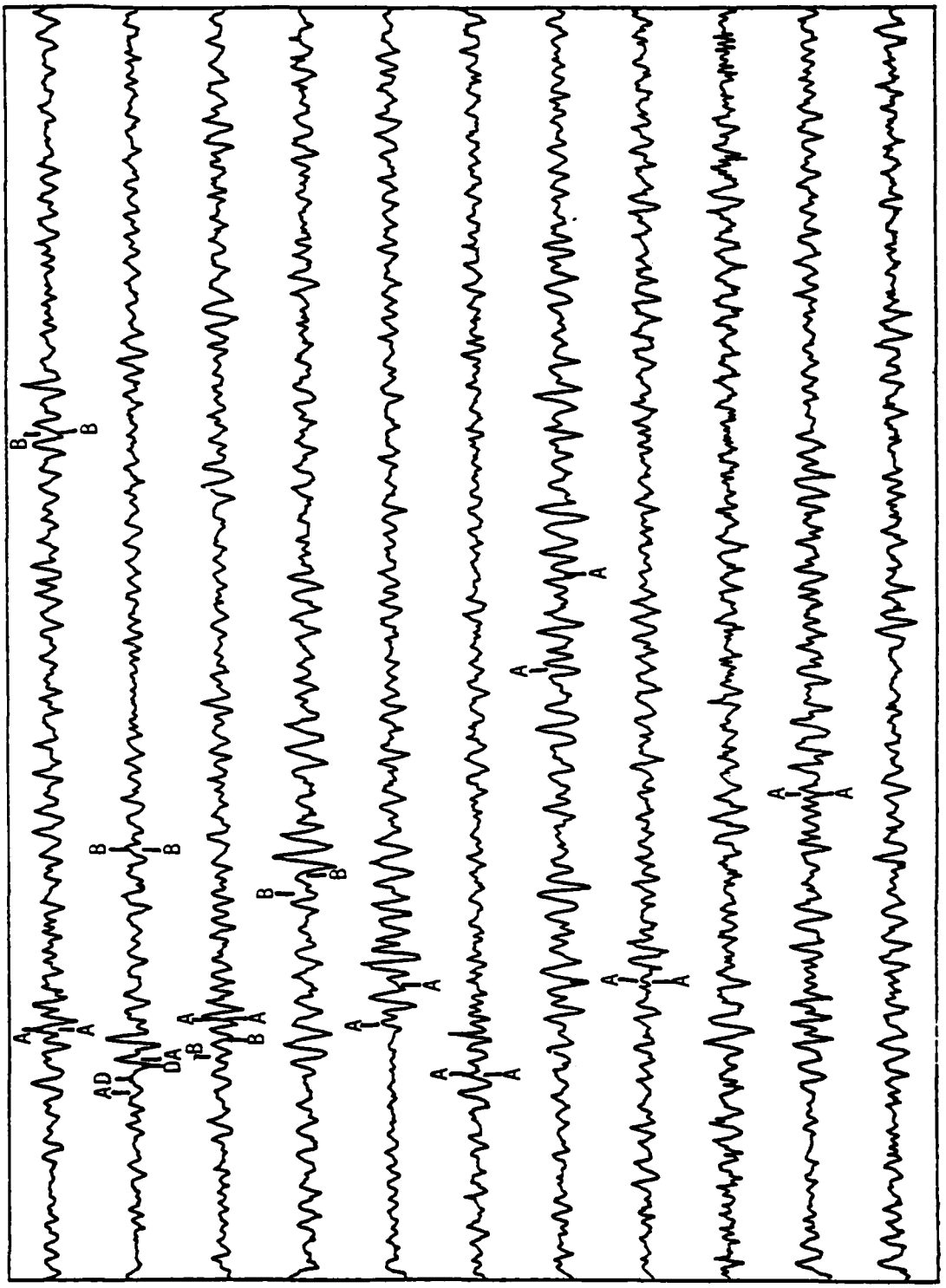

Figure 7b.

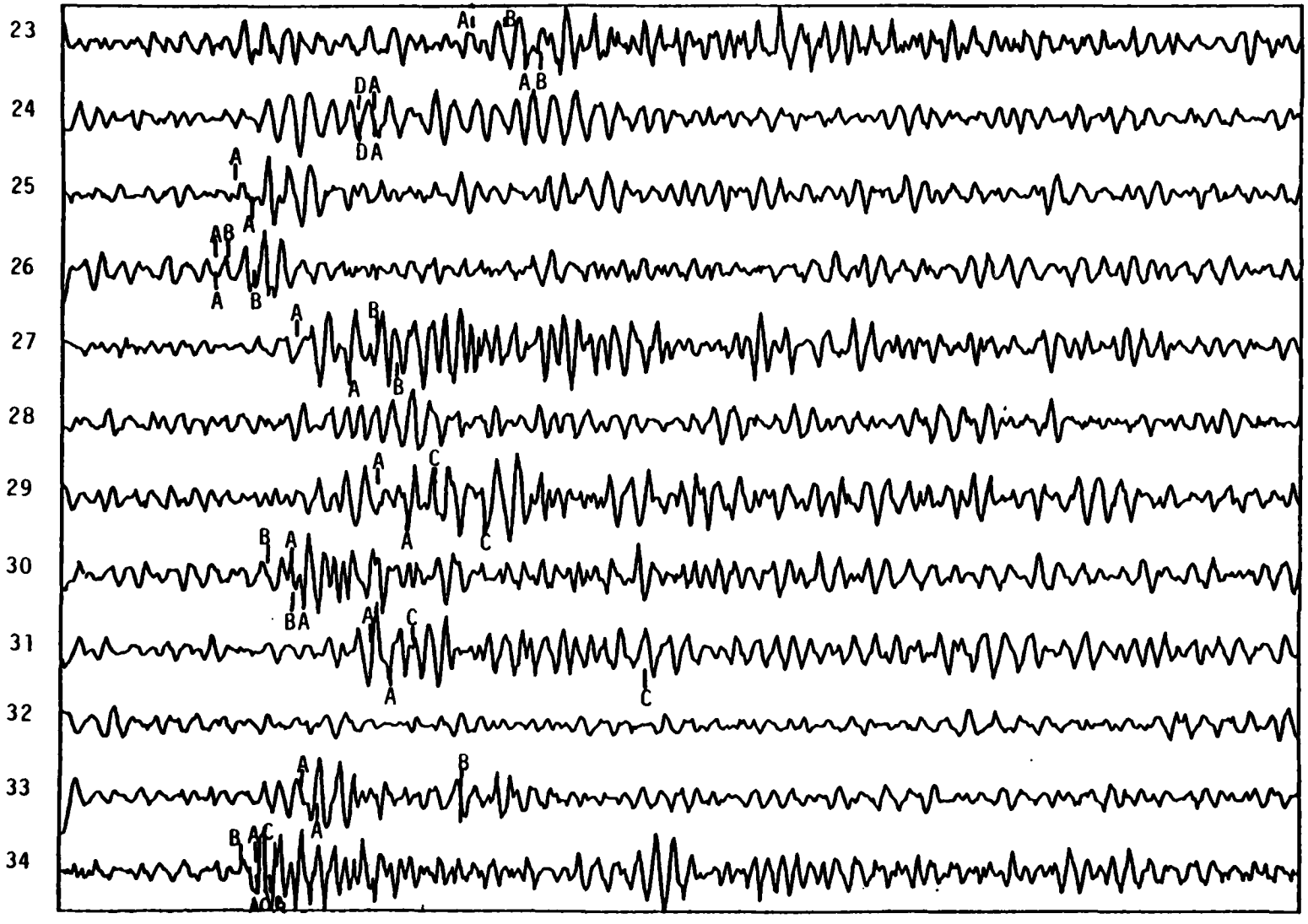

Figure 7c.

Figure 8a. Detection results for the Pinedale test tape. Letters A through D show the amplitude level of the signal for which the labeled phase was detected. Symbols above each trace apply to earliest picks; symbols below each trace apply to picks with the largest mean relative amplitude (MRA). Overall false alarm data as well as a summary of detections are presented in Table 2.

Figure 8b.

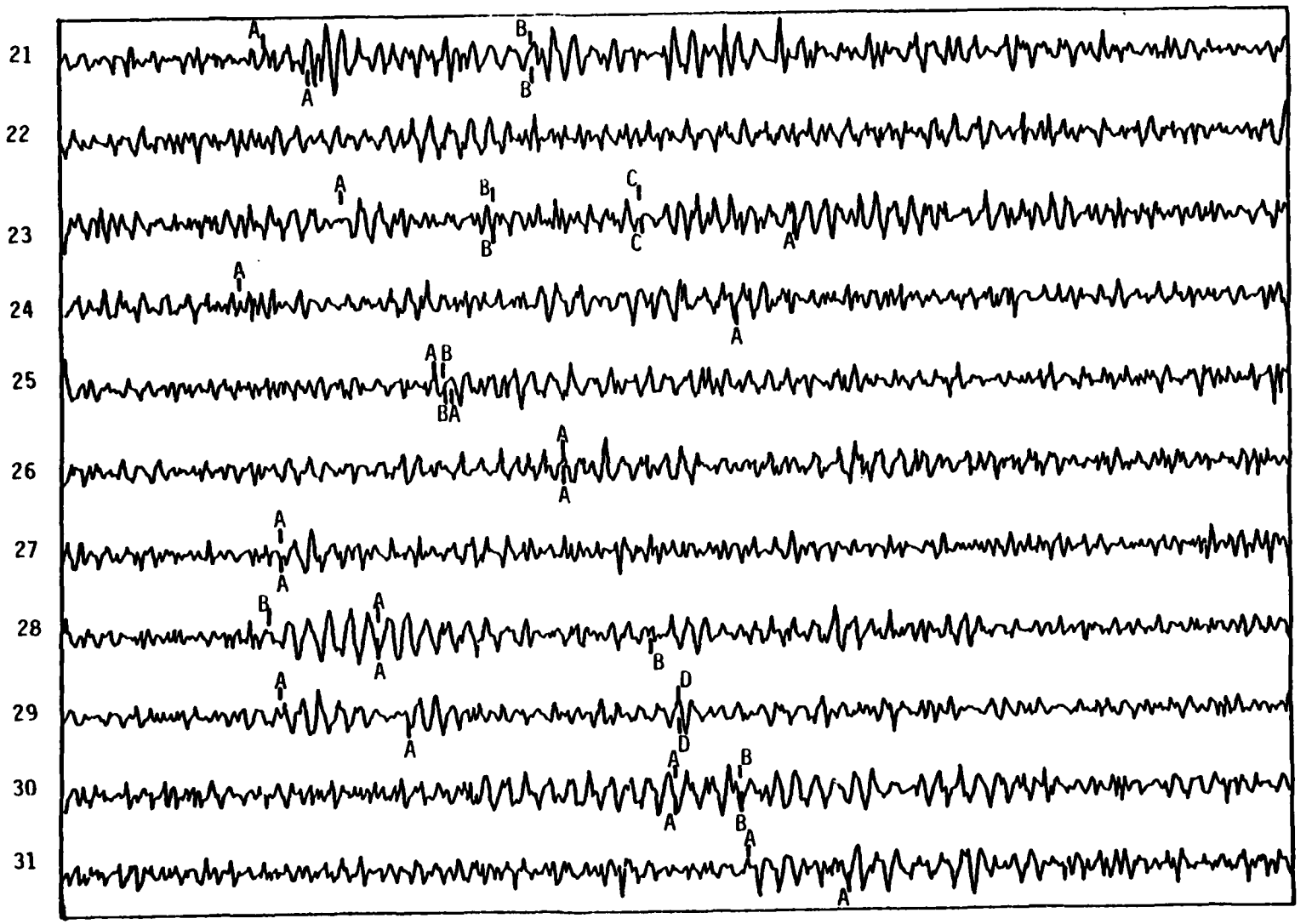

Figure 8c.

TABLE 1

SUMMARY LISTING OF DETECTION RESULTS FOR THE NORSAR TAPE USING THE MARS ALGORITHM

S^3 Class	A		B		C		D			
Raw SNR	1/2		1/4		1/8		1/16			
	FA	D	FA	D	FA	D	FA	D	ΣFA	ΣD
Event										
1	2	1	0	1	0	1	1	0	3	3
2	1	1	0	1	0	1	0	0	1	3
3	0	1	0	1	0	1	1	1	1	4
4	0	0	1	0	1	0	1	0	3	0
5	1	1	1	1	2	1	0	0	4	3
6	0	0	1	0	1	0	0	0	2	0
7	0	0	0	0	1	0	1	0	2	0
8	1	1	0	1	1	0	0	0	2	2
9	0	1	0	0	2	0	0	0	2	1
10	0	1	2	0	1	0	1	0	4	1
11	1	1	1	1	0	0	1	0	3	2
12	1	1	2	1	0	0	1	0	4	2
13	1	1	0	1	1	0	0	1	2	3
14	0	1	0	1	2	0	2	0	4	2
15	1	0	1	1	1	0	1	0	4	1
16	0	1	2	0	1	0	1	0	4	1
17	1	1	0	0	1	0	1	0	3	1
18	0	1	1	0	1	0	0	0	2	1
19	2	1	0	0	0	0	2	0	4	1
20	2	0	1	0	3	0	1	0	7	0
21	1	1	1	0	2	0	2	0	6	1
22	0	0	2	0	3	0	1	0	6	0
23	0	1	0	1	0	0	0	0	0	2
24	0	1	1	0	2	0	1	1	4	2
25	1	1	1	0	0	0	2	0	4	1
26	1	1	1	1	1	0	0	0	3	2
27	2	1	1	1	0	0	2	0	5	2
28	1	0	0	0	0	0	0	0	0	1
29	0	1	0	0	1	1	1	0	2	2
30	0	1	0	1	2	0	1	0	3	2
31	1	1	0	0	1	1	1	0	3	2
32	1	0	1	0	1	0	0	0	3	0
33	0	1	2	1	0	0	0	0	2	2
34	0	1	2	1	0	1	1	0	3	3
ΣD		26		16		7		3		52
ΣFA	22		25		32		27		106	
Percent D		79		48		21		9		

TABLE 2

SUMMARY LISTING OF DETECTION RESULTS FOR THE NORSAR TEST TAPE
FROM THE VSC BENCHMARK ANALYSIS (BLANDFORD, PERSONNEL COMMUNICATION)

S^3 Class	A		B		C		D			
Raw SNR	1/2		1/4		1/8		1/16			
	FA	D	FA	D	FA	D	FA	D	ΣFA	ΣD
Event										
1	3	1	1	1	2	1	2	0	8	3
2	2	1	1	1	1	1	1	0	5	3
3	2	1	1	1	1	1	1	0	5	3
4	2	1	1	0	2	0	2	0	7	1
5	0	1	1	1	1	0	2	0	4	2
6	2	0	0	0	2	0	3	0	7	0
7	3	0	1	0	2	0	2	0	8	0
8	4	1	2	1	1	0	1	0	8	2
9	1	1	1	0	0	0	0	0	2	1
10	2	1	0	0	1	0	0	0	3	1
11	1	1	2	1	0	0	0	0	3	2
12	1	1	2	0	0	0	3	0	6	1
13	3	1	2	0	1	0	3	0	9	1
14	2	1	1	1	1	0	0	0	4	2
15	3	1	2	0	2	0	2	0	9	1
16	1	1	1	1	2	0	0	0	4	2
17	0	1	1	0	1	0	2	0	4	1
18	0	1	1	0	1	0	0	0	2	1
19	0	0	1	0	2	0	3	0	6	0
20	1	0	3	1	1	0	1	1	6	2
21	1	1	0	1	0	0	2	0	3	2
22	2	1	0	0	0	1	1	0	3	2
23	0	1	2	1	1	0	0	0	3	2
24	0	1	0	0	0	0	0	1	0	2
25	0	1	0	0	2	0	3	0	5	1
26	0	1	0	1	2	0	0	0	2	2
27	2	1	0	0	3	0	1	0	6	1
28	1	0	1	0	2	0	0	0	4	0
29	1	1	2	1	4	0	2	0	9	2
30	1	1	1	1	2	0	1	0	5	2
31	1	1	0	0	2	0	1	0	4	1
32	1	0	1	0	1	0	2	0	5	0
33	0	1	0	1	0	0	0	0	0	2
34	0	1	3	1	2	0	1	0	6	2
ΣD		28		16		4		2		50
ΣFA	43		35		45		42		165	
Percent D		80		50		10		6		

detections missed in the STA/LTA run. MARS missed eight events declared in the STA/LTA run, but 3 of these were of questionable quality in the STA/LTA run.

The false alarm rate (FAR) of the MARS run is significantly below that of the STA/LTA run: The STA/LTA detector had approximately 7.1 false alarms per hour, compared with 4.7 false alarms per hour for MARS. In summary, MARS yielded 13 percent more detections with a FAR 30 percent less than the STA/LTA detector.

4.2 PINEDALE

The Pinedale detection test tape is similar to the NORSAR tape in that known signals are buried in phase-randomized seismic noise. The data differs in that (a) the digitization rate is 20 Hz (NORSAR was 10 Hz decimated from 20 Hz) and (b) the 31 signals are generally longer and have lower dominant frequencies.

Using the same criteria as described for the NORSAR test, MARS found 53 good detections out of a total of 173 declared events, resulting in a FAR of 5.8/hour (Figure 8, Table 3). In comparison, the STA/LTA detector had 47 detections and a FAR of 5.6/hour (Table 4). Again, MARS compares quite favorably with the STA/LTA detector, locating 13 percent more events for essentially the same FAR.

Of the 31 Class A events, MARS detected all but one, Event 22. The Pinedale events generally have a narrower bandwidth and a more emergent onset when compared to the NORSAR data. Thus, although MARS is able to detect a fair percentage of the signals, the resulting time picks often occur well after the nominal nine-minute event "arrival" time. In many cases, an analyst would have considerable difficulty calling the first arrival at the ninth minute also; therefore, we have followed the previous convention of declaring a pick to be valid if it is correlated with a recognizable phase in the corresponding Class A signal.

4.3 COMPARISONS WITH THE VSC BENCHMARK①

Analog presentations of the results from the VSC benchmark runs are not available, so it is impossible to make comparisons between the two studies on an event by event basis. Perusal of Figures 7 and 8 show some obvious failures by the MARS detector (NORSAR events 4A and 28A; Pinedale event 22A). Pinedale event 16 (see Figure 8b) is a convincing success. We have adopted a somewhat more rigourous criterion for accepting picks within the signal coda, but then have admitted a few early picks, primarily because it seems that many of the signals start well before the nominal nine minute mark. The problem of accurately identifying "first motion" remains.

A phenomenon seen in both the VSC and MARS results is the occasional event which is missed at one amplitude level but then found at a weaker one. This is not an impossible occurrence, and does not necessarily mean that the low level detection is wrong. Using elementary probability theory, we can estimate the probabilities for all possible pairs of signal detection results. Consider the Pinedale Class B and Class C events. If we let $p(B)$ and $p(C)$ be the probabilities of a B event and a C event being separately found, then Table 4 shows $p(B) = 0.39$ and $p(C) = 0.29$. Let $B \cdot C$ denote an event detected at both amplitude levels, $\bar{B} \cdot \bar{C}$ an event missed at both amplitude levels, and similarly for $B \cdot \bar{C}$ and $C \cdot \bar{B}$. Then the joint events should have the following probabilities

$$p(B \cdot C) = p(B)p(C)$$

$$p(\bar{B} \cdot \bar{C}) = (1-p(B))(1-p(C))$$

$$p(\bar{B} \cdot C) = (1-p(B))p(C)$$

$$p(B \cdot \bar{C}) = (1-p(C))p(B)$$

Table 5 shows that observed joint occurrences are not too different from expectation.

The MARS detector clearly exceeded the VSC benchmark for the Class C signals, suggesting appreciably better performance at low

① Later work at VSC (Blandford et al., 1981) using fixed bandwidth filters (1.5 - 4 Hz for NORSAR and 1 - 2 Hz for Pinedale) produced STA/LTH results not appreciably different from the MARS RESULTS.

TABLE 3

SUMMARY LISTING OF DETECTION RESULTS FOR THE PINEDALE TEST TAPE USING THE MARS ALGORITHM

S^3 Class	A		B		C		D			
Raw SNR	1/2		1/4		1/8		1/16			
	FA	D	FA	D	FA	D	FA	D	ΣFA	ΣD
Event										
1	1	1	1	0	1	0	0	0	3	1
2	0	1	0	1	1	0	1	0	2	2
3	0	1	0	1	1	1	2	0	3	3
4	0	1	1	0	1	0	0	0	2	1
5	0	1	0	0	1	0	1	0	2	1
6	0	1	3	0	2	0	1	0	6	1
7	1	1	2	0	0	1	2	0	5	2
8	1	1	1	0	3	0	1	0	6	1
9	2	1	2	0	1	0	2	0	7	1
10	2	1	2	0	1	0	0	0	5	1
11	0	1	0	1	1	0	0	0	1	2
12	0	1	0	0	0	1	0	0	0	2
13	0	1	0	1	2	1	1	0	3	3
14	1	1	0	0	1	0	1	0	3	1
15	0	1	1	1	1	1	0	0	2	3
16	0	1	0	1	0	1	0	0	0	3
17	1	1	0	1	1	1	1	0	3	3
18	1	1	2	0	2	1	1	0	6	2
19	1	1	1	0	0	0	2	0	4	1
20	1	1	0	1	2	0	1	1	4	3
21	2	1	1	1	0	0	1	0	4	2
22	1	0	1	0	0	0	3	0	5	0
23	2	1	3	1	2	1	1	0	8	3
24	1	1	4	0	1	0	0	0	6	1
25	0	1	1	0	1	0	1	1	3	2
26	1	1	1	0	2	0	1	0	5	1
27	1	1	0	0	1	0	1	0	3	1
28	0	1	1	1	0	0	1	0	2	1
29	1	1	1	0	1	0	1	1	4	2
30	2	1	1	1	3	0	1	0	7	2
31	0	1	2	0	3	0	1	0	6	1
ΣD		30		12		9		2		53
ΣFA	23		32		36		29		120	
Percent D		97		39		29		6		

TABLE 4

SUMMARY LISTING OF DETECTION RESULTS FOR THE PINEDALE TEST TAPE
FROM THE VSC BENCHMARK ANALYSES (BLANDFORD PERSONNEL COMMUNICATION)

S^3 Class	A		B		C		D			
Raw SNR	1/2		1/4		1/8		1/16			
	FA	D	FA	D	FA	D	FA	D	ΣFA	ΣD
Event										
1	2	1	1	0	2	0	1	0	6	1
2	0	1	1	1	1	0	0	0	2	2
3	0	1	1	1	0	1	0	0	1	3
4	1	1	3	0	2	0	0	0	6	1
5	2	0	1	0	1	0	0	0	4	0
6	0	1	1	1	2	0	0	0	3	2
7	0	1	2	1	0	0	2	0	4	2
8	0	1	3	0	1	0	2	0	6	1
9	0	1	0	0	2	0	1	1	3	2
10	0	1	2	1	2	0	0	1	4	3
11	0	1	0	1	0	0	2	0	2	2
12	0	1	3	0	1	0	3	0	7	1
13	0	1	0	1	3	1	1	0	4	3
14	1	0	1	0	0	0	0	0	2	0
15	1	1	0	0	1	0	0	0	2	1
16	1	1	1	1	5	0	0	0	7	2
17	0	1	0	1	2	0	0	0	2	2
18	1	1	0	0	2	0	0	0	3	1
19	0	1	0	0	1	0	1	0	2	1
20	1	1	1	0	3	0	0	0	5	1
21	3	1	0	1	1	1	0	0	4	3
22	0	1	1	1	0	0	0	0	1	2
23	0	1	0	1	1	0	2	0	3	2
24	0	1	0	0	1	0	1	0	2	1
25	1	1	2	0	2	0	1	0	6	1
26	0	0	0	0	1	0	2	0	3	0
27	0	1	0	0	2	0	1	0	3	1
28	1	1	0	0	1	0	1	0	3	1
29	1	1	1	0	0	0	1	0	3	1
30	0	1	1	1	2	0	3	0	6	2
31	0	1	1	1	1	0	3	0	5	2
ΣD		28		14		3		2		47
ΣFA	16		27		43		28		114	
Percent D		90		45		10		6		

signal levels. The reason for this is not known, but may be related to the somewhat longer equivalent short term integration time. The MARS filter comb was perhaps a little too coarse, and some improvement might accrue if the filter had more frequency overlap. The filter impulse response time was nearly ideal for the one- and two-second teleseismic body waves, but should be made ten times longer to pick up low amplitude regional phases.

Another respect in which the MARS results seem better than the VSC benchmark is the variability of the false alarm rate. This is especially true for the Pinedale tape.

To summarize, the MARS detector is about as fast as the VSC detector (when allowance is made for the different computers), performs significantly better on weak signals and appears to have a more regular false alarm behavior.

TABLE 5

JOINT PROBABILITIES FOR PINEDALE CLASS B AND CLASS C EVENTS

	$p(\overline{B} \cdot \overline{C})$	$p(B \cdot \overline{C})$	$p(\overline{B} \cdot C)$	$p(B \cdot C)$
Observation	0.51	0.19	0.10	0.19
Theory	0.44	0.26	0.17	0.11

V. FUTURE MARS RESEARCH

The MARS detection results reported above were obtained using a fairly simple algorithm for recognizing collections of peaks. The current algorithm simply searches for a minimum number of peaks above a certain threshold within a fixed-width time window in the t_g-f plane. Even so, the present MARS detection capability is close to that predicted for an ideal matched filter operating on a completely known signal (see Appendix 2). It is thus unlikely that the present level of detection capability can be greatly exceeded by any detector. However, it may be possible to lower the false alarm rate by implementation of a more sophisticated test procedure. For example, the spectral continuity of a true signal can be utilized: the spectrum of a signal varies slowly with frequency, so its peaks should occur in adjacent frequency bands. A random line-up of noise peaks, however, tends to be scattered in frequency. Thus one test may be to require the adjacency of a fraction of the peaks in a collection before an event is declared. The size and shape of the time window is also a discriminating factor. A narrower window that allows some dispersion (such a window would be slanted in the t_g-f plane, i.e., later at high frequencies) encompasses less area in the t_g-f plane, thus reducing the chance that noise peaks will form a valid collection.

An appreciable fraction of the work reported here has been concerned with signal processing -- implementation of efficient Gaussian filters, pattern recognition of peak collections, etc. However, future research will have to be more concerned with seismology and probabilistic analysis of factors affecting the false alarm rate. Of prime importance should be a study of the characteristics that distinguish seismic events from seismic noise in the t_g-f plane. Shown in Figure 9 are plots of the t_g-f plane for Pinedale Events 16A, 16B and 16C (the same signal recorded at three different levels). These plots (termed "sonograms" in acoustics) have proven useful for examining the time/frequency distribution of signals and noise.

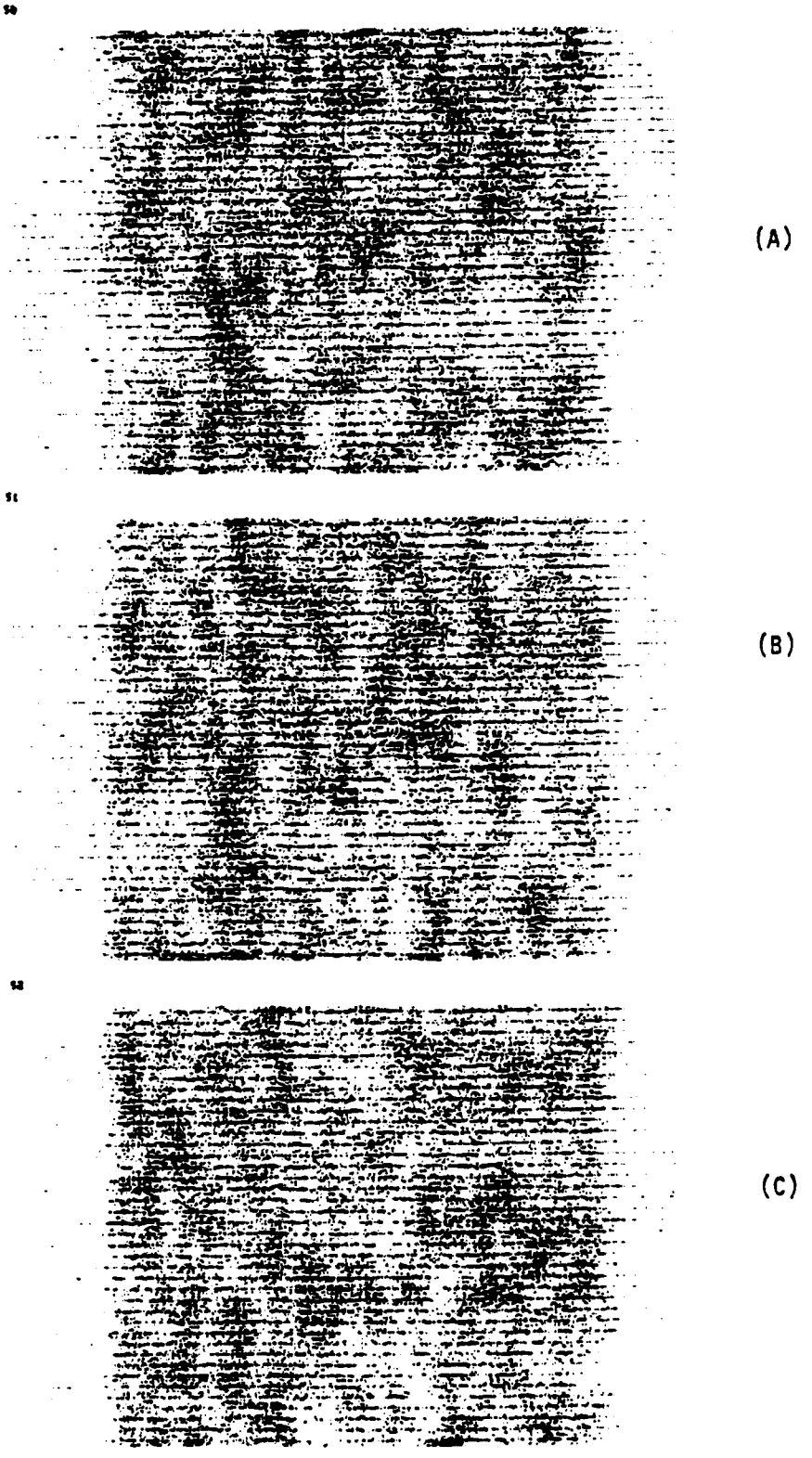

Figure 9. "Sonograms" for event 16 from the Pinedale test tape (see Figure 6b). The phase is clearly present in panels A and B, and can be located in panel C, as well, where it was found by the detector despite the apparent marginal signal-to-noise ratio.

VI. CONCLUSIONS

The MARS seismic event detector offers a significant improvement over the current VSC optimally-filtered STA/LTA detector. On nearly 45 hours of synthetic data, MARS detected 13 percent more events than the STA/LTA detector, demonstrating its capability of extracting low-level signals in a poor SNR environment. The additional events detected by MARS are nearly all small amplitude Class B and C events. The advantage of the MARS algorithm at low signal levels should not be surprising, since MARS is not simply a power-law detector but also uses the signal dispersion and bandwidth as discriminating characteristics.

The improvement in detections was achieved with no attendant increase in the false alarm rate. In fact, with the NORSAR data, the MARS FAR was only two-thirds that of the STA/LTA detector. The MARS FAR with the Pinedale data is equal to that of the STA/LTA detector. It is likely that the MARS FAR can be lowered with the implementation of more discriminating detection tests, but the present level of MARS detections is close to the theoretical limit predicted for an ideal matched filter and is unlikely to be greatly increased by any means.

The ten percent advantage in probability of detection shown by the MARS detector for this particular class of events is equivalent to an improvement of about 0.1 in body wave magnitude. That is, the MARS detector should have nearly the same probability of detection and same false alarm rate as the benchmark VSC detector, but for signals which are an average of 0.1 magnitude smaller.

To make this calculation it is necessary to estimate the derivative of the curve relating probability of detection to signal-to-noise ratio at constant false alarm rate. Since the receiver operating characteristics for the seismic detection algorithms are unknown, we take as a model the problem of the detection of a completely known signal in white noise. The relevant curves have been plotted by Helstrom (1968, Figure IV.1) and are

reproduced in Figure A2.1 in Appendix A2 to this report. Reading from the curve for false alarm probability $Q_0 = 10^{-2}$, a variation in detection probability from 80 percent to 90 percent is equivalent to an increase in signal-to-noise ratio from 3.0 to 3.5. The differential body wave magnitude is the logarithm of the ratio of these two numbers, $\Delta m_b = -0.07$. Since the detector operating characteristic is nonlinear, the exact value for the body wave magnitude differential depends on the reference signal-to-noise ratio about which the variation is taken; for weaker signals (hence lower probability of detection), the magnitude differential increases.

This is, of course, only a very rough calculation. To make a better estimate would require the Monte Carlo estimation of the exact receiver operating characteristic for each detection algorithm and signal type.

In general, it is felt that the detection experiment was wisely run. The range of signal amplitudes perfectly straddled the range of interest. The phase scrambling seems effectively to have smeared out short term fluctuations in the noise. Several snapshot power spectra calculated for arbitrary noise windows showed, as might be expected, only small variation in the noise background. They were much less than are expected to occur over longer periods of time, so the sensitivity of the FAR to seasonal variations in the noise background has not satisfactorily been explored. It is also difficult to quantify the real probability of detection because a variety of signal types were included. Tests with one signal, at one amplitude but buried 10 or so times in different noise realizations would help here. We are also not convinced that it is necessary to use actual recorded ground noise for hiding the signals. By all the statistical measurements we have made, it is difficult if not impossible to distinguish earth noise from random numbers with a suitably shaped autocorrelation function. Since there are now machine-independent techniques for generating pseudo-random data, the detection experiment could have been

conducted by distributing a random-number program and a spectral shaping program, using the data tape just for seismic signals. This would, we feel, facilitate some of the additional measurements proposed above.

REFERENCES

Berger, J. (1980), "Seismic Detectors: The State of the Art," Systems, Science and Software Report SSS-R-80-4588.

Blandford, R., D. Racine and R. Romaine (1981), "Single Channel Event Detection," Teledyne-Geotech, (in press).

Helstrom, C. W. (1968), "Statistical Theory of Signal Detection," Pergamon Press, New York.

Papoulis, A. (1965), "Probability, Random Variables and Stochastic Processes," McGraw-Hill, New York.

APPENDIX 1
EFFECT OF WINDOW LENGTH ON FALSE ALARM RATE

Assume that the false alarms (called events) follow a Poisson model, so that the longer the time that has elapsed from the previous event, the more likely another one is to occur. For this model, the probability that exactly k events happen within a window t seconds in length is (Papoulis, 1965, p. 558)

$$p_k(t) = \frac{(\lambda t)^k}{k!} \exp(-\lambda t) \qquad \text{A1.1}$$

hence, the probability of no events is

$$p_0(t) = \exp(-\lambda t) \qquad \text{A1.2}$$

and the probability of at least one event is

$$1-p_0(t) = 1 - \exp(-\lambda t) \qquad \text{A1.3}$$

Suppose for t = 100 sec, there are 100 events (false alarms) in 24 hours of data. Then

$$1-p_0(100) = \frac{100}{864} = 0.1157 \qquad \text{A1.4}$$

and A1.3 yields

$$\lambda = 1.23 \times 10^{-3} \qquad \text{A1.5}$$

Using A1.3 again, if we decrease the window length to 2 seconds, the probability of at least one event in any given window becomes

$$1-p_0(2) = 2.6 \times 10^{-3} \qquad \text{A1.6}$$

Since there are 43,200 two-second windows in 24 hours, there will be a total of 112 false alarms. This is only a twelve percent increase in false alarm rate.

APPENDIX 2.
ESTIMATION OF SIGNAL-TO-NOISE RATIO FOR BROADBAND SEISMIC SIGNALS

A common seismological measure of signal-to-noise ratio (SNR) is the quantity

$$\frac{\text{maximum peak-to-peak signal amplitude}}{\text{2 times rms noise}} \qquad \text{A2.1}$$

This is the quantity sometimes referred to as "raw signal to noise ratio." In attempting to relate the empirical results obtained in the seismic detection experiment to fundamental theoretical results in the field of statistical detection theory, it quickly becomes apparent that although the quantity defined in A2.1 may be useful for the relative ranking of signals, it is a very unsatisfactory absolute measure. In order to connect more closely the behavior of seismic event detectors to the various statistical models, it is important to derive a better seismological measure of SNR. The reason for this is that as deterministic methods of waveform modeling continually improve, one asks at each stage what are the consequences for seismic event identification.

Most present seismic detectors are in one way or another related to incoherent, or power-law, detectors, and make the weakest possible assumptions about the signal being sought. Suppose there were available a near-perfect replica of a signal known to be hidden in earth noise. Would this knowledge have important consequences for identifying and discriminating events?

At first sight the answer might seem to be yes, but the issue is not at all clear. For example, when one compares the theoretical result for the probability of detecting a completely known signal in white noise to the result when the amplitude spectrum, but not the phase spectrum is known, for false alarm probabilities of 10^{-2} and signal-to-noise ratio 4, the probability of detection is 92 percent in the former case and 90 percent in the latter (Helstrom, 1968, Figure IV.1 and Figure V.2).

Stated differently, for the same false alarm rate and 90 percent detection probability, the loss incurred in not knowing the phase spectrum is only 1.5 db (equivalent to 0.08 magnitude units).

The simplest theoretical problem is the detection of a completely known signal in white noise. The "best" (in the least squares sense) detector for this case is the correlation detector or matched filter. There is no theorem which proves a better detector to be impossible to construct, but it is the ideal reference or standard by which other detectors, including seismic detectors, can be judged.

The performance of the matched filter detector is presented in Figure A2.1. This figure depicts, for three different false alarm probabilities (10^{-2}, 10^{-3}, 10^{-4}), the connection between signal-to-noise ratio and probability of detection. The signal-to-noise ratio in this figure is defined as

$$d = \sqrt{2E/N} \qquad \text{A2.2}$$

where N is the one-sided power spectral density of the noise (assumed white and Gaussian) and the signal energy E is defined by

$$E = \int_0^T [s(t)]^2 dt \qquad \text{A2.3}$$

with s(t) the time signal, and T its duration in time.

The difficulty in using the quantity defined in A2.1 as a measure of the signal-to-noise ratio becomes apparent when the results obtained from the seismic detection experiment are compared to the data in Figure A2.1. For example, it was found for the A class events (raw SNR = 1/2, by definition A2.1) that the detection probability was about 90 percent, for a false alarm rate of about 5/hour.

Five false alarms each hour corresponds to a false alarm rate, Q_o, of about 10^{-3}. Suppose one assumes that d = 1/2 and Q_o = 10^{-3} (point A on the plot). Reading the abscissa for point A, it

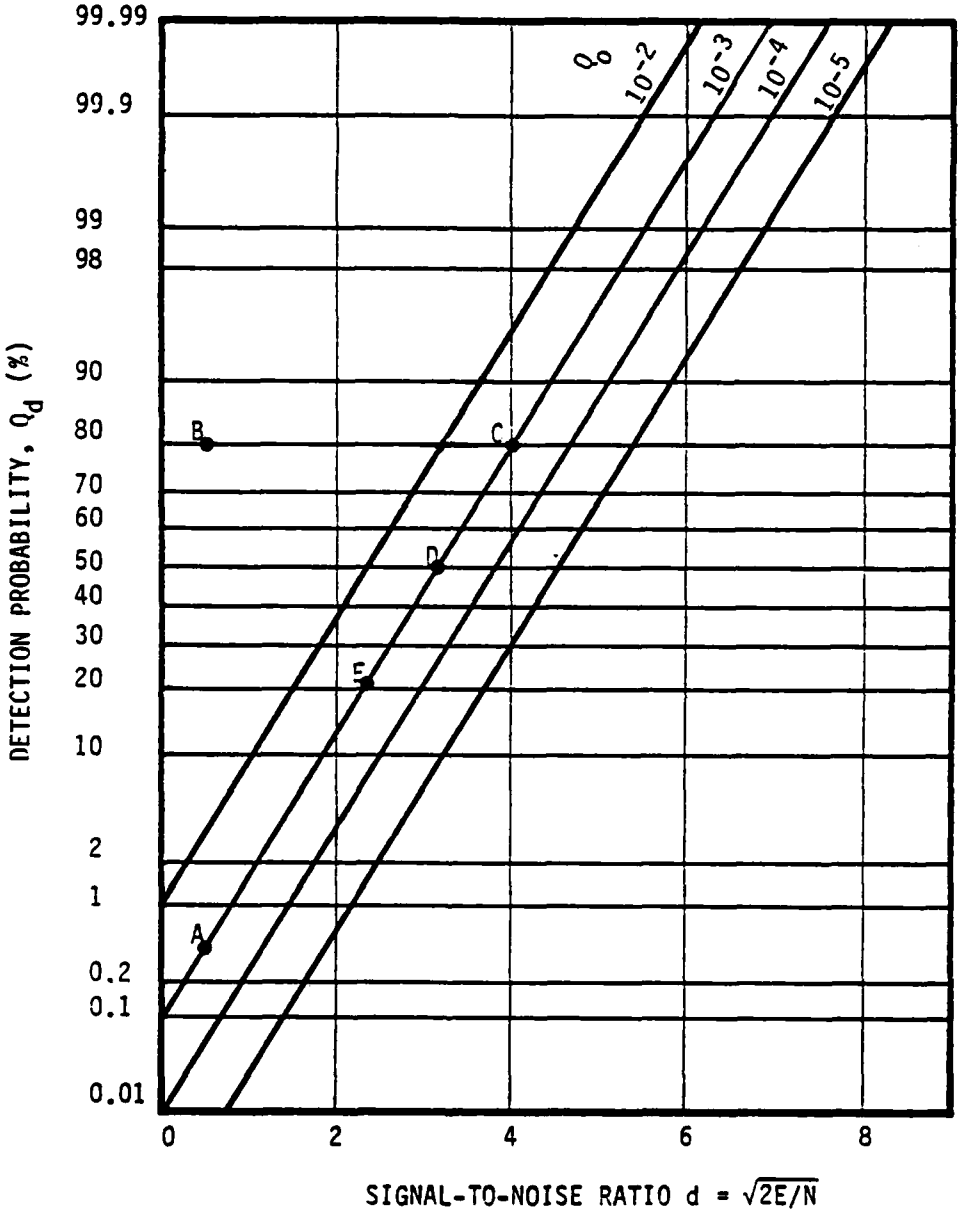

Figure A2.1. Probability of detecting a completely known signal, Q_d, plotted against Q_o the false alarm probability with parameter d the signal-to-noise ratio.

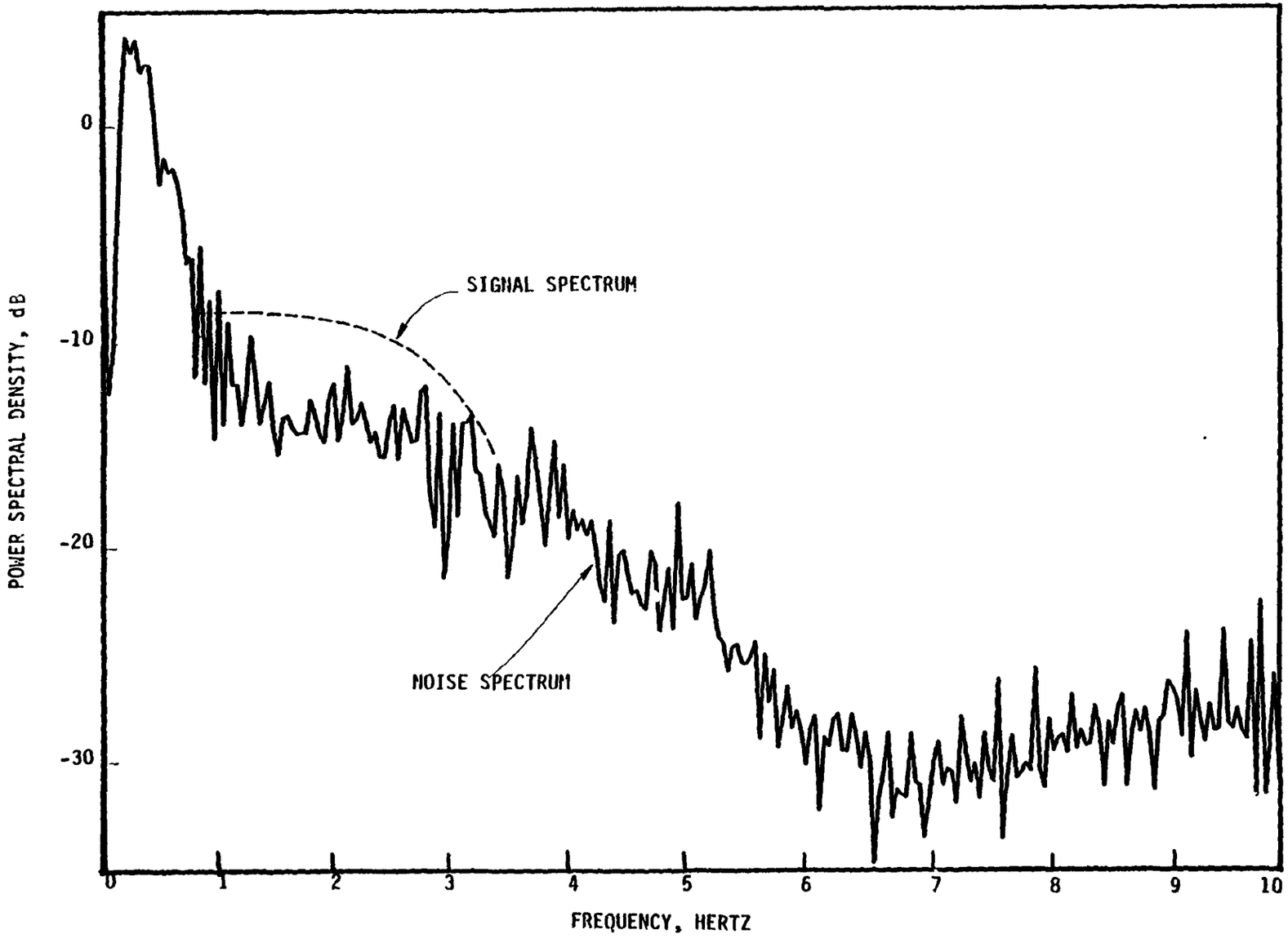

Figure A2.2 Power spectrum calculated for a sample of Pinedale noise, with schematic representation of a typical signal spectrum.

is seen that the matched filter would have a detection probability below one percent, but the observed detection probability (point A) is nearer 90 percent. In view of the inhomogeneous character of the seismic signals, the radical variation in the noise spectrum with frequency, and the simple nature of the detectors used, it is implausible to believe that the seismologists have made a detector orders of magnitude more sensitive than the matched filter.

The alternative way to enter the figure is to accept that $Q_0 = 10^{-3}$, and that the detection probability equals 80 percent (point C). This implies that the matched filter with this performance would require a signal-to-noise ratio, d, of about 4. Presumably the seismic detectors would need even greater values.

Further evidence that equation A2.1 underestimates the true signal-to-noise ratio by a factor of about ten is obtained by comparing the results of the detector experiment for the three largest raw signal-to-noise ratios (1/2, 1/4, 1/8). Typical detection probabilities (Table 1, for instance) are 0.8, 0.5, and 0.2. These points are labeled B, C, and D on Figure A1. Reading the abscissas, we see that a matched filter would require the signal-to-noise ratio, d, to be 4, 3, and 2 respectively. Not exactly the four-to-one range present in the data, but roughly consistent with the known relative amplitudes.

It is easy to see that the reason definition A2.1 vastly underestimates the actual signal-to-noise ratio is because it uses much too large a value for the noise factor in the denominator. Consider the earth noise power spectrum displayed in Figure A2.2. This particular spectrum was calculated over a noise window in the Pinedale data tape, but the NORSAR noise spectra are similar, except for the smaller Nyquist frequency. The predominant feature in this spectrum, and in spectra calculated for data in many other sites with similar instrument response functions, is the knee around 1 Hz and the steep rise in power at lower frequencies. The dashed curve superimposed on the spectrum shows the general character of weak but detectable earthquake and explosion seismic signals. They emerge

from the low frequency noise hump at about 1 Hz, slowly bend downwards and disappear once again into the noise background at some higher frequency. (The signal spectrum shown here, which falls below the noise spectrum for frequencies greater than 3 Hz, is typical of many of the spectra on the Pinedale tape. Other events at other stations (NORSAR) can contain energy to 5 Hz and beyond). When a time domain calculation such as that defined by equation A2.1 is used to estimate the signal-to-noise ratios, it is incapable of recognizing that the signal energy and noise energy have quite different frequency distribution. In fact, turning again to Figure A2.2, suppose the low frequency hump were four times as high, but that the spectrum above 1 Hz were unchanged. This is not an artificial example, because the hump is due to microseisms which do vary from day to day. If the hump were twice as high, the rms noise amplitude would be twice as large. If the signal were unchanged, it would seem by definition A2.1 that the signal-to-noise ratio had halved. Yet, it seems plausible (and can be mathematically established) that more or less arbitrary amounts of noise power can be added to a signal with no degradation in detectibility, as long as the frequency distribution of the noise power and the frequency distribution of the signal power do not overlap.

It can be seen, therefore, that definition A2.1 invariably underestimates the actual signal-to-noise ratio. We prefer to accept the more pessimistic definition of signal-to-noise ratio: "The most realistic measure of signal-to-noise ratio is the one which is largest."

Having berated the usual seismological definition of signal-to-noise ratio, we now show how a practical calculation very much like A2.1 can be devised which is much closer to generally accepted theoretical definitions.

The starting point is the definition given by Helstrom, Equation IV, 3.19

$$d^2 = \int_{-\infty}^{\infty} \frac{|S(f)|^2}{\phi(f)}\, df \qquad \text{A2.4}$$

where S is the Fourier Transform of the signal (dimension amplitude x time), ϕ is the power spectral density of the noise (dimension amplitude squared x time), and d is dimensionless. Turning again to Figure A2.2, it is clear that the predominant contribution to the integral defined by equation A2.4 comes from the frequency interval $0.8 < f < 3.0$ Hz. Furthermore, it can be seen that the power spectral density over this band is nearly flat and, taking -12.5 db as the average, has the value

$$N = 2\Phi(f) = 0.056 \qquad \text{A2.5}$$

(We follow Helstrom using Φ for the two-sided power spectral density function and N for the one-sided and constant spectral density function.)

With the assumption that the noise spectrum is <u>locally</u> white, Equations A2.4 and A2.5 lead to the estimate

$$d^2 = \frac{2}{N} \int_{-\infty}^{\infty} |S(f)|^2\, df \qquad \text{A2.6}$$

Using Parsevals theorem,

$$d^2 = \frac{2}{N} \int_{-\infty}^{\infty} (s(t))^2\, dt \qquad \text{A2.7}$$

and since the integral of the squared time domain amplitude is just the signal energy, we have arrived back at Equation A2.2

It is now an easy step to formulate the procedure for calculating signal-to-noise ratios from time domain records in the case where the noise is approximately white over the band in which the signal energy is contained. The first procedure is to filter the broad band seismogram to eliminate those parts of the spectrum

where $|S(f)|^2/\phi(f) < 1$. In the present case, this is easy, for it is only necessary to high-pass the data to eliminate the low frequency noise hump. The spectrum plotted in Figure A2.2 indicates that the corner frequency for this high-pass filter should be located at about 1.0 Hz.

Because of the steep edge on the high frequency side of the noise hump, the exact value of the corner frequency should be picked with care. Several filters were applied to the NORSAR and Pinedale Class A signals, and the best visual results were obtained for f_c = 0.8 Hz. Deviations as small as 0.1 Hz either way produced noticibly poorer results. In these data, as with most seismic data, a low pass operation is not so critical, for the signal falls back below the noise at about the frequency at which the noise spectrum itself decays due to the combined effect of instrument response function and anti-alias filters. Thus, the high frequency noise power is considerably less than the noise power within the signal band, even though the total bandwidth is much greater. The Class A signals displayed in Figures 7 and 8 were treated in this way.

The next step is to choose a time window T_s which exactly contains the signal. Several other time windows (the exact length need not be the same) are taken over sections of the record without signal. For each window, it is next necessary to calculate the variance of the data according to

$$\text{var}(x) = \frac{1}{T}\int_0^T (x(t))^2\, dt \qquad \text{A2.8}$$

The square root of the variance is the standard deviation (SD) or RMS amplitude. Table A2.1 shows RMS amplitudes for two signal windows and five noise windows, each 1.5 seconds long taken from two events for the Pinedale data plotted in Figure 7. The parenthetical values in the signal window column are signal standard deviation corrected for the effect of noise in the usual manner

TABLE A2.1

RMS SIGNAL AMPLITUDES (1.5 SEC WINDOW) FROM PINEDALE TEST TAPE

Event	Seismic Window	Noise Windows		
2	0.76 (0.63)	0.34	0.39	0.52
6	0.77 (0.64)	0.35	0.38	

$$\text{var(total)} = \text{var(signal)} + \text{var(noise)} \qquad \text{A2.9}$$

Since signal energy is the product of the variance and the time duration, we have

$$\sqrt{E} = \sqrt{T_s}\ SD(S) \qquad \text{A2.10}$$

Furthermore, a good estimate of the noise power spectral density when it is approximately white is

$$N = \frac{\text{var}(n)}{B_n} \qquad \text{A2.11}$$

where B_n is an estimate of the bandwidth. Equation A2.11 follows from the fundamental definition that the integral of a power spectrum over all frequencies is just the variance in the time series. Although the choice of B_n is somewhat arbitrary, it is never as large as the Nyquist frequency, $f_N = 1/2\Delta t$. Putting A2.10 and A2.11 into A2.2 we have

$$d \approx \sqrt{2T_s B_n}\ \ \frac{SD(s)}{SD(n)} \qquad \text{A2.12}$$

The quantity defined by Equation A2.12 is closely allied to the cruder estimates given by Equation A2.1. The principal distinction is the dimensionless time-bandwidth product, but also important is the pre-filter operation to limit the noise bandwidth to coincide with the known signal bandwidth.

The correction mentioned above turns out to be significant. Using the data presented in Table A2.1, and taking $B_n = 2.2$ Hz (3.0 to 0.8) we find

$$d = 4.0 \qquad \text{A2.13}$$

These signal-to-noise ratios are some eight times larger than those obtained from definition A2.1, and bring the observed detector performance results more in accord with the predictions of the matched filter processor. If anything, the numbers derived from

Equation A2.12 are probably still too small because they lie uncomfortably near to the equivalent matched filter characteristic indicated by point C on Figure A2.1.

An alternate, and somewhat more reliable way to estimate N is from the power spectrum itself. Using the value for the (flat) power spectral density given in Equation A2.5, we get

$$d = 4.5 \qquad \text{A2.14}$$

The effect of the time duration factor, T_s, contained in Equation A2.11 can be important for drawn out regional phases such as Lg. Event 3 in the NORSAR data (Figure 5) lasts for nearly 40 seconds. If the bandwidth and power level of the noise are about the same as those previously quoted for the Pinedale data, and if the signal bandwidth is not too different, the signal-to-noise ratio for this record ought to be $\sqrt{40/1.5}$ greater than for the short teleseismic P-wave of equal RMS amplitude. Thus $d_{LG} \approx 21$ suggesting that a strategy radically different from STA/LTA or the present short time duration MARS process should be used for detecting regional phases.

The frequency-time decomposition of the instantaneous signal energy intrinsic to the MARS analysis permits a more refined estimation of the signal-to-noise ratio. Although this calculation has not been implemented, it would be a straightforward extension to the algorithm. The proposed method is a discrete approximation to Equation A2.4

$$d^2 = \int_{-\infty}^{\infty} \frac{|S(f)|^2}{\phi(f)}\, df \qquad \text{A2.4}$$

with allowance made for the fact that the signal energy may not arrive at the same time for different frequency components.

To estimate signal-to-noise ratio for non-stationary seismic signals, recall first that the suite of narrowband envelope

functions for a time series x, $E_x(f,t)$, is an estimate of the instantaneous Fourier Spectrum of x. Let x(f) be a section of record consisting of noise alone, taken shortly before the signal arrives. Then we can estimate the noise power spectral density by

$$\phi(f) = A < E_n(f,t) > /\Delta f \qquad \text{A2.15}$$

where < > denotes an average in time, A is a normalization constant and Δf is the filter bandwidth. Now take a later time window which straddles the signal, and let $\Delta t(f)$ be the frequency dependent time window containing the signal as well, of course, as the background noise. Assume that the noise is stationary from the start of the noise window to the end of the signal window. Then an estimate of the signal spectrum is

$$\left| S(f) \right| = B \left| <E_s(f,t)>_{\Delta t(f)} - <E_n(f,t)> \right| \qquad \text{A2.16}$$

where B is another normalization constant and $< >_{\Delta t(f)}$ denotes an average over the signal time window. This equation states that the best estimate of the signal amplitude spectrum at each frequency is the sum of the instantaneous envelope amplitudes, reduced by the best estimate of the noise amplitudes. The noise correction is required to ensure that if $\left|S(f)\right|$ is evaluated over a segment with no signal present, its expected value is zero.

Equation A2.15 and A2.16 can be combined in the manner implied by Equation A2.4 to yield

$$d^2_{MARS} = \sum_{f=f_{min}}^{f_{max}} C \left| \frac{<E_s>^2_{\Delta t(f)}}{<E_n>^2} - 1 \right| \qquad \text{A2.17}$$

wnere C is a new constant which depends on frequency (possibly) and filter bandwidth, and where we neglect the cross product between signal spectrum and noise spectrum which arises when Equation A2.16 is squared.